

Small-time fluctuations for the bridge of a sub-Riemannian diffusion

Ismael Bailleul¹,
Laurent Mesnager² & James Norris^{3,4}

May 14, 2015

Abstract

We consider small-time asymptotics for diffusion processes conditioned by their initial and final positions, under the assumption that the diffusivity has a sub-Riemannian structure, not necessarily of constant rank. We show that, when the endpoints are joined by a unique path of minimal energy, the conditioned diffusion converges weakly to that path. We show further that, when the endpoints lie outside the sub-Riemannian cut locus, the fluctuations of the conditioned diffusion from the minimal energy path, suitably rescaled, converge to a Gaussian limit. The Gaussian limit is characterized in terms of the minimal second variation of the energy functional at the minimal path, the formulation of which is new in this context.

1 Introduction

Let M be a connected C^∞ manifold of dimension d and let a be a C^∞ non-negative quadratic form on the cotangent space T^*M . We assume that a has a sub-Riemannian structure, that is to say, there exist $m \in \mathbb{N}$ and C^∞ vector fields X_1, \dots, X_m on M such that

$$a(\xi, \xi) = \langle \xi, a(x)\xi \rangle = \sum_{\ell=1}^m \langle \xi, X_\ell(x) \rangle^2, \quad \xi \in T_x^*M \quad (1)$$

and such that

$$\text{span}\{Y(x) : Y \in \mathcal{A}(X_1, \dots, X_m)\} = T_x M, \quad \text{for all } x \in M. \quad (2)$$

¹Institut de Recherche Mathématiques de Rennes, 263 Avenue du General Leclerc, 35042 Rennes, France

²Modal'x, Université de Paris X, 200 avenue de la République, 92001 Nanterre, France

³Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WB, UK

⁴Research supported by EPSRC grant EP/103372X/1

Here $\mathcal{A}(X_1, \dots, X_m)$ denotes the smallest set of vector fields on M containing X_1, \dots, X_m and closed under the commutator product, given by

$$[X, Y](x) = \sum_{i=1}^d X^i(x) \frac{\partial Y}{\partial x^i}(x) - Y^i(x) \frac{\partial X}{\partial x^i}(x).$$

There is associated to the quadratic form a an energy function I on the set of continuous paths $\Omega = C([0, 1], M)$. An absolutely continuous path $\omega \in \Omega$ may have a driving path ξ , by which we mean a measurable path ξ in T^*M such that $\xi_t \in T_{\omega_t}^*M$ and $\dot{\omega}_t = a(\omega_t)\xi_t$ for almost all t . Then ω has energy

$$I(\omega) = \int_0^1 \langle \xi_t, a(\omega_t)\xi_t \rangle dt.$$

If ω is not absolutely continuous or has no driving path, then $I(\omega) = \infty$. Write H for the subset of Ω consisting of paths of finite energy. For $x, y \in M$, set

$$H^{x,y} = \{\omega \in H : \omega_0 = x, \omega_1 = y\}.$$

It is well known that $H^{x,y}$ is non-empty under the bracket condition (2) and that the sub-Riemannian distance

$$d(x, y) = \inf_{\omega \in H^{x,y}} \sqrt{I(\omega)} \tag{3}$$

defines a metric compatible with the topology of M .

In this paper we investigate the restriction of the energy function I to $H^{x,y}$ and its relation to heat flow from x to y . We make no assumption of completeness for M so in general the infimum in (3) may not be achieved. Nevertheless, we will show in Theorem 1.1 that the small-time logarithmic asymptotics for heat kernels are determined by the distance function and indeed that the associated diffusion measures concentrate exponentially fast in small time around paths of nearly minimal energy.

The main work in the paper addresses the case where x and y are chosen so that I achieves a minimum on $H^{x,y}$ uniquely, say at γ . We will then construct, under a regularity condition on γ , a vector space $T_\gamma H^{x,y}$ of absolutely continuous paths v in TM , with $v_t \in T_{\gamma_t}M$ for all t and $v_0 = v_1 = 0$, along with an equivalence class of norms on $T_\gamma H^{x,y}$, each making $T_\gamma H^{x,y}$ into a Hilbert space. The paths in $T_\gamma H^{x,y}$ can be thought of as the infinitesimal variations of γ in $H^{x,y}$. We will further construct a continuous non-negative quadratic form Q on $T_\gamma H^{x,y}$ such that $Q(v)$ is the minimal second variation of I in the direction v , in a sense to be made precise. These constructions are the content of Section 7.

The sub-Riemannian cut locus was defined by Bismut [7] in terms of the bicharacteristic flow associated to a . An alternative characterization was then given by Ben Arous [6] in terms of a quadratic form on control paths associated to a compatible sub-Riemannian structure. We will show in Theorem 8.3 that (x, y) lies outside the sub-Riemannian cut locus if and only if our

regularity condition on γ holds and Q is positive-definite. Since the vector space $T_\gamma H^{x,y}$ and the quadratic form Q are intrinsic to a , this provides an intrinsic characterization of the cut locus in terms of the energy function in this general setting. To clarify the distinction between objects intrinsic to a and those constructed from a compatible sub-Riemannian structure, we show by example in Section 7 that, when a has non-constant rank, it may admit two inequivalent sub-Riemannian structures. For control theory, the sub-Riemannian structure is the basic given quantity, but for heat flow it is natural to seek objects intrinsic to a .

The sub-Riemannian cut locus is less well understood than its Riemannian counterpart. It is known to be a closed and symmetric subset of $M \times M$. Rifford & Trélat [27] and Agrachev [2] have proved results which limit its size. The reader may find in the lecture notes of Agrachev, Barilari & Boscain [1, Theorems 10.4 and 10.11] a proof that, for all $x \in M$ and any $r \in (0, \infty)$ such that $B = \{y \in M : d(x, y) \leq r\}$ is compact, the set of points $y \in B$ such that (x, y) lies outside the cut locus is dense in B .

Let \mathcal{L} be a second order differential operator on M with C^∞ coefficients, such that $\mathcal{L}1 = 0$ and such that \mathcal{L} has principal symbol $a/2$. In each coordinate chart, \mathcal{L} thus takes the form

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^d a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^d b^i(x) \frac{\partial}{\partial x^i} \quad (4)$$

for some C^∞ functions b^i on M . In this context we refer to a as the diffusivity. Given that a has sub-Riemannian structure (X_1, \dots, X_m) , we obtain the same family of operators by choosing another C^∞ vector field X_0 on M and setting⁵

$$\mathcal{L} = \frac{1}{2} \sum_{\ell=1}^m X_\ell^2 + X_0. \quad (5)$$

We will sometimes impose the following global condition

$$M = \mathbb{R}^d \quad \text{and} \quad X_0, X_1, \dots, X_m \in C_b^\infty(\mathbb{R}^d, \mathbb{R}^d). \quad (6)$$

At other times, we will impose no further condition on M or a , in particular no condition of completeness, but then we will insist that there is a C^∞ 1-form β satisfying the sector condition

$$\|a(\beta, \beta)\|_\infty < \infty \quad (7)$$

and there is a locally invariant⁶ positive C^∞ measure ν on M such that

$$\mathcal{L}f = \frac{1}{2} \operatorname{div}(a \nabla f) + a(\beta, \nabla f) \quad (8)$$

⁵Here, X_ℓ is used to denote the differential operator $\sum_{i=1}^d X_\ell^i(x) \partial / \partial x^i$.

⁶When M is compact, the local invariance property does not restrict the class of operators. To see this, suppose that \mathcal{L} has the form (8) for some positive C^∞ measure ν , not necessarily invariant. Then \mathcal{L} has a

where the divergence is understood with respect to ν . Here, by a positive C^∞ measure ν we mean any measure having a positive C^∞ density with respect to Lebesgue measure in each coordinate chart. We say that ν is locally invariant for \mathcal{L} if, for all C^∞ functions f of compact support in M , we have

$$\int_M a(\beta, \nabla f) d\nu = 0. \quad (9)$$

The divergence of a vector field X with respect to ν is characterized by

$$\int_M f \operatorname{div} X d\nu = - \int_M X f d\nu$$

for all such functions f . From (8), we find that $X_0 = \sum_{\ell=1}^m \alpha_\ell X_\ell$, where $\alpha_\ell = \frac{1}{2} \operatorname{div} X_\ell + \langle \beta, X_\ell \rangle$. In particular, $X_0(x) \in \operatorname{span}\{X_1(x), \dots, X_m(x)\}$ for all $x \in M$, which is not required in (6). The sector condition and the local invariance property are satisfied trivially in the case $\beta = 0$. Thus our results apply to all sub-Riemannian Laplacians.

There is a family of probability measures on the set $\Omega^{x,y} = \{\omega \in \Omega : \omega_0 = x, \omega_1 = y\}$ which is naturally associated to the operator \mathcal{L} . Fix $\varepsilon > 0$ and $x \in M$. There exists a diffusion process starting from x and having generator $\varepsilon \mathcal{L}$. Since, in general, the coefficients of \mathcal{L} may be unbounded and we make no assumption of completeness for M , this diffusion may explode with positive probability, that is to say it may leave all compact sets in finite time. We will write μ_ε^x for the unique sub-probability measure on Ω which is the law of this diffusion restricted to paths which do not explode by time 1. Under our assumptions, there is a unique family of probability measures $(\mu_\varepsilon^{x,y} : y \in M)$ on Ω which is weakly continuous in y , with $\mu_\varepsilon^{x,y}$ supported on $\Omega^{x,y}$ for all y , and such that

$$\mu_\varepsilon^x(d\omega) = \int_M \mu_\varepsilon^{x,y}(d\omega) p(\varepsilon, x, dy)$$

where $p(\varepsilon, x, \cdot)$ is the (sub-)law of ω_1 under μ_ε^x . More explicitly, the finite-dimensional distributions of each measure $\mu_\varepsilon^{x,y}$ may be written as follows. Choose a positive C^∞ measure ν on M . It does not have to be the same measure as in the preceding paragraph. There is a positive C^∞ function p on $(0, \infty) \times M \times M$ such that

$$p(\varepsilon, x, dy) = p(\varepsilon, x, y) \nu(dy).$$

positive C^∞ invariant density function ρ with respect to ν on M . Thus, for all C^∞ functions f , we have.

$$\int_M a(\beta \rho - \frac{1}{2} \nabla \rho, \nabla f) d\nu = \int_M \mathcal{L} f \rho d\nu = 0.$$

Set $\tilde{\beta} = \beta - \frac{1}{2} \nabla \log \rho$ and $\tilde{\nu} = \rho \nu$. Then $\mathcal{L} f = \frac{1}{2} \widetilde{\operatorname{div}}(a \nabla f) + a(\tilde{\beta}, \nabla f)$, $\tilde{\beta}$ satisfies (7), and $\tilde{\nu}$ is invariant. In particular, when M is compact and a is positive-definite, we can write all operators \mathcal{L} of the form (4) in the form (8).

This function p is the Dirichlet heat kernel for \mathcal{L} with respect to ν . Then, for all $k \in \mathbb{N}$, all $t_1, \dots, t_k \in (0, 1)$ with $t_1 < t_2 < \dots < t_k$ and all $x_1, \dots, x_k \in M$, we have

$$\begin{aligned} \mu_\varepsilon^{x,y}(\{\omega : \omega_{t_1} \in dx_1, \dots, \omega_{t_k} \in dx_k\}) \\ = \frac{p(\varepsilon t_1, x, x_1)p(\varepsilon(t_2 - t_1), x_1, x_2) \dots p(\varepsilon(1 - t_k), x_k, y)}{p(\varepsilon, x, y)} \nu(dx_1) \dots \nu(dx_k). \end{aligned}$$

It is straightforward to see that these finite-dimensional distributions are consistent and do not depend on the choice of ν .

Our main aim is to understand the behaviour of the diffusion bridge measures $\mu_\varepsilon^{x,y}$ in the limit $\varepsilon \rightarrow 0$. We will say that $\gamma \in H^{x,y}$ is strongly minimal if there exist $\delta > 0$ and a relatively compact open set U in M such that

$$I(\omega) \geq I(\gamma) \text{ for all } \omega \in H^{x,y} \quad \text{and} \quad I(\omega) \geq I(\gamma) + \delta \text{ for all } \omega \in H^{x,y} \text{ which leave } U.$$

When M is complete for the sub-Riemannian distance, the second condition follows from the first by the Hopf–Rinow theorem. Moreover, if γ is the only strongly minimal path in $H^{x,y}$, then the second condition holds for all relatively compact domains U containing γ by a weak compactness argument. Here is our first main result.

Theorem 1.1. *Let M be a connected C^∞ manifold. Let \mathcal{L} be a second order differential operator on M of the form (8),*

$$\mathcal{L}f = \frac{1}{2} \operatorname{div}(a \nabla f) + a(\beta, \nabla f)$$

where the diffusivity a has a sub-Riemannian structure, where the divergence is taken with respect to a positive C^∞ locally invariant measure, and where β is a C^∞ 1-form satisfying the sector condition (7). Then, for all $x, y \in M$, as $t \rightarrow 0$, we have

$$t \log p(t, x, y) \rightarrow -d(x, y)^2/2. \quad (10)$$

Moreover, for all $\delta > 0$, for $r = \delta^{1/4}(d(x, y)^2 + \delta)^{1/2}$, we have

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon^{x,y}(\{\omega \in \Omega^{x,y} : d(\omega_t, \Gamma_t(\delta)) \geq r \text{ for some } t \in [0, 1]\}) \leq -\delta/2 \quad (11)$$

where

$$\Gamma_t(\delta) = \{\gamma_t : \gamma \in H^{x,y}, I(\gamma) \leq d(x, y)^2 + \delta\}.$$

In particular, in the case where there is a unique strongly minimal path $\gamma \in H^{x,y}$, we have $\mu_\varepsilon^{x,y} \rightarrow \delta_\gamma$ as $\varepsilon \rightarrow 0$ weakly on $\Omega^{x,y}$, where δ_γ is the unit mass at γ .

The small-time logarithmic asymptotics (10) were proved by Varadhan [29] in the case when $M = \mathbb{R}^d$ and a is everywhere positive-definite. They were generalized by Léandre [17, 18] to the sub-Riemannian case, under the global assumption (6), using probabilistic techniques which make this assumption unavoidable. While the lower bound extends easily, for the upper bound it appears that analytic techniques are needed, and we arrive at (10) by checking the applicability of a general result of Sturm [28] in this context. In the case of Brownian motion in a complete Riemannian manifold, the exponential estimate (11) follows from a result of Hsu [11]. Under the assumption (6) and subject to the condition that $a(x)$ is positive-definite, (11) follows from work of Inahama [13].

We now go beyond the law of large numbers and exponential asymptotics contained in Theorem 1.1 to identify a small-time equivalent for the heat kernel and to obtain a central limit theorem for the diffusion bridge measures. For this we will assume that x and y are chosen so that

$$\text{there is a unique strongly minimal path } \gamma \in H^{x,y} \quad (12)$$

and further that (x, y) lies outside the sub-Riemannian cut locus. We first state a small-time equivalent for the heat kernel, which was proved by Ben Arous [6] under the condition (6). In fact, Ben Arous obtained a full asymptotic expansion, with local regularity statements, but we will be content with the basic form.

Theorem 1.2. *Under the conditions of Theorem 1.1, in the case where there is a unique strongly minimal path $\gamma \in H^{x,y}$ and (x, y) lies outside the cut locus, there is a constant $c(x, y) > 0$ such that, in the limit $t \rightarrow 0$,*

$$p(t, x, y) = c(x, y)t^{-d/2} \exp\{-d(x, y)^2/(2t)\}(1 + o(1)).$$

We turn to the second order result which is the true focus of this work. The quadratic form Q mentioned above allows us to give $T_\gamma H^{x,y}$ an intrinsic Hilbert norm. Write $T_\gamma \Omega^{x,y}$ for the set of continuous paths v in TM such that $v_t \in T_{\gamma_t} M$ for all t and $v_0 = v_1 = 0$. We make $T_\gamma \Omega^{x,y}$ into a Banach space using the uniform norm $\|v\|_\infty = \sup_{t \in [0,1]} |v_t|$ corresponding to a choice of Riemannian metric on M . The associated topology on $T_\gamma \Omega^{x,y}$, which is all that matters for us, does not depend on the choice of metric. We will show in Theorem 8.5 that there is a unique zero-mean Gaussian measure μ_γ on $T_\gamma \Omega^{x,y}$ such that

$$\int_{T_\gamma \Omega^{x,y}} \phi(v)^2 \mu_\gamma(dv) = Q(\tilde{\phi})$$

for all continuous linear functionals ϕ on $T_\gamma \Omega^{x,y}$, where $\tilde{\phi} \in T_\gamma H^{x,y}$ is determined by $\phi(v) = Q(\tilde{\phi}, v)$ for all $v \in T_\gamma H^{x,y}$.

We rescale the fluctuations of the diffusion bridge around the minimal path γ to obtain a non-degenerate limit. To do this, we choose a C^∞ map $\theta : M \rightarrow \mathbb{R}^d$ such that the derivative

$\theta^*(\gamma_t) : T_{\gamma_t}M \rightarrow \mathbb{R}^d$ is invertible for all $t \in [0, 1]$. We can always choose a chart U along γ and obtain a suitable function θ by extending the coordinate map from a neighbourhood of γ . Define $\sigma_\varepsilon : \Omega^{x,y} \rightarrow T_\gamma\Omega^{x,y}$ by

$$\sigma_\varepsilon(\omega)_t = \theta^*(\gamma_t)^{-1}(\theta(\omega_t) - \theta(\gamma_t))/\sqrt{\varepsilon}.$$

Then we obtain a probability measure $\tilde{\mu}_\varepsilon^{x,y}$ on $T_\gamma\Omega^{x,y}$ by setting

$$\tilde{\mu}_\varepsilon^{x,y} = \mu_\varepsilon^{x,y} \circ \sigma_\varepsilon^{-1}.$$

It is straightforward to check that if $\tilde{\mu}_\varepsilon^{x,y}$ converges weakly on $T_\gamma\Omega^{x,y}$ as $\varepsilon \rightarrow 0$ for one choice of the function θ , then it does so for all such choices and with the same limit.

Theorem 1.3. *Let M be a connected C^∞ manifold. Let \mathcal{L} be a second order differential operator on M , which has the form (4) in each coordinate chart, and whose diffusivity has a sub-Riemannian structure. Assume that either (5) and (6) hold or (7) and (8) hold. Let $x, y \in M$. Suppose that there is a unique strongly minimal path $\gamma \in H^{x,y}$ and that (x, y) lies outside the cut locus. Then the rescaled diffusion bridge measures satisfy $\tilde{\mu}_\varepsilon^{x,y} \rightarrow \mu_\gamma$ weakly on $T_\gamma\Omega^{x,y}$ as $\varepsilon \rightarrow 0$.*

In the case of Brownian motion on a compact Riemannian manifold, the convergence $\tilde{\mu}_\varepsilon^{x,y} \rightarrow \mu_\gamma$, with μ_γ characterized by its covariance, was derived by Molchanov [22], but a full proof was not given. Many of the techniques we use follow ideas pioneered by Bismut [7] and Ben Arous [6] in their studies of the heat kernel, indeed, it turns out that the diffusion bridge asymptotics are another side of the same story. Two recent works have taken the fundamental ideas of Bismut and Ben Arous in different directions to ours. First, Deuschel, Friz, Jacquier & Violante [9, 10] have obtained small time expansions for marginal distributions of the heat flow. Second, Barilari, Boscain & Neel [5] have found estimates of the heat kernel actually on the cut locus.

In Section 2, we recall the notions of the bicharacteristic flow associated to a and the sub-Riemannian cut locus. We also give a characterization of the limit fluctuation measure μ_γ in terms of the bicharacteristic flow. Then, in Section 3, we discuss the case where $a(x)$ is positive-definite for all x , and thus defines a Riemannian metric. In this case the limit measure μ_γ has several different characterizations using classical objects of Riemannian geometry. The proof of Theorem 1.3 is given under the global assumption (6) in Section 4, adapting a method of Azencott, Bismut and Ben Arous, and relying on ideas of Malliavin calculus. Then in Section 5, we show how to localize to a neighbourhood of the minimal path, and we prove Theorem 1.2 and complete the proof of Theorem 1.3 using analytic estimates in Section 6. Section 6 also contains the proof of Theorem 1.1 and does not depend on material in other sections. As outlined above, Sections 7 and 8 are devoted to the geometry of some spaces of paths, leading to the intrinsic construction of the minimal second variation Q of the energy

function and its associated Gaussian measure. These two sections may be read independently of the preceding probabilistic and analytic parts.

We are grateful to Emmanuel Trélat and Fabrice Baudoin for helpful advice in the course of this work.

2 Cut locus and fluctuation measure in terms of the bicharacteristic flow

The bicharacteristic flow is the maximal flow $(\psi_t(\lambda) : \lambda \in T^*M, t \in (\zeta^-(\lambda), \zeta^+(\lambda)))$ of the vector field V on T^*M given by $\beta(V, \cdot) = d\mathcal{H}$, where β is the canonical symplectic 2-form on T^*M and $\mathcal{H} : T^*M \rightarrow [0, \infty)$ is the Hamiltonian

$$\mathcal{H}(\lambda) = \frac{1}{2} \langle \lambda, a(x)\lambda \rangle, \quad \lambda \in T_x^*M.$$

Thus, for all $\lambda \in T^*M$, we have $\psi_0(\lambda) = \lambda$ and $\zeta^-(\lambda) < 0 < \zeta^+(\lambda)$, and

$$\dot{\psi}_t(\lambda) = V(\psi_t(\lambda)), \quad t \in (\zeta^-(\lambda), \zeta^+(\lambda))$$

and $\psi_t(\lambda)$ leaves all compact sets in T^*M as $t \rightarrow \zeta^+(\lambda)$ if $\zeta^+(\lambda) < \infty$ and as $t \rightarrow \zeta^-(\lambda)$ if $\zeta^-(\lambda) > -\infty$. If M is complete, then $\zeta^\pm(\lambda) = \pm\infty$ for all $\lambda \in T^*M$. In a coordinate chart U for M , for $\lambda \in U$, write $\lambda_t = \psi_t(\lambda) = (x_t, p_t)$ while λ_t remains in T^*U , with $x_t \in U$ and $p_t \in \mathbb{R}^d$. Then, for any sub-Riemannian structure (X_1, \dots, X_m) for a , we have

$$\begin{aligned} \dot{x}_t &= \sum_{\ell=1}^m \langle p_t, X_\ell(x_t) \rangle X_\ell(x_t), \quad x_0 = x \\ \dot{p}_t &= - \sum_{\ell=1}^m \langle p_t, X_\ell(x_t) \rangle \langle p_t, \nabla X_\ell(x_t) \rangle, \quad p_0 = p \end{aligned} \tag{13}$$

where $\langle p, \nabla X(x) \rangle_i = (\partial/\partial x^i) \langle p, X(x) \rangle$.

Fix $x, y \in M$ and suppose that condition (12) holds, that is to say, there is a unique strongly minimal path $\gamma \in H^{x,y}$. Write π for the projection $T^*M \rightarrow M$. We assume that γ is a normal minimizer, meaning that

$$\text{there exists a bicharacteristic } (\lambda_t)_{t \in [0,1]} \text{ such that } \gamma_t = \pi \lambda_t \text{ for all } t \in [0,1]. \tag{14}$$

For $t \in [0,1]$, write $\lambda_t = \psi_t(\lambda_0)$ and define linear maps $J_t : T_x^*M \rightarrow T_{\gamma_t}^*M$ and $K_t : T_y^*M \rightarrow T_{\gamma_t}^*M$ by

$$J_t \xi_0 = \left. \frac{\partial}{\partial \varepsilon} \right|_{\varepsilon=0} \pi \psi_t(\lambda_0 + \varepsilon \xi_0), \quad K_t \xi_1 = \left. \frac{\partial}{\partial \varepsilon} \right|_{\varepsilon=0} \pi \psi_{-(1-t)}(\lambda_1 - \varepsilon \xi_1). \tag{15}$$

We assume that (x, y) is non-conjugate for $(\lambda_t)_{t \in [0,1]}$, meaning that

$$J_1 \text{ is invertible.} \quad (16)$$

Following Bismut [7], when conditions (12), (14) and (16) hold, the pair (x, y) is said to lie outside the cut locus of a . By a simple and well known argument, (16) implies that the bicharacteristic projecting to γ is unique.

The following statement is part of Theorem 8.5.

Theorem 2.1. *Assume the hypotheses of Theorem 1.3. Then the limit measure μ_γ is the unique zero-mean Gaussian measure on $T_\gamma \Omega^{x,y}$ with covariance given for $s \leq t$ by*

$$\int_{T_\gamma \Omega^{x,y}} v_s \otimes v_t \mu_\gamma(dv) = J_s J_t^{-1} K_t^*. \quad (17)$$

The characterization of the cut locus and limit measure μ_γ in terms of the bicharacteristic flow is computationally effective and does not require the construction of the quadratic form Q . On the other hand, the formulation of Theorem 1.3 in terms of Q confirms in the setting of an infinite-dimensional path space an intuition derived from analogous considerations for functions and measures in finite dimensions. Moreover, the proof of Theorem 1.3 is by analysis in path space and leads naturally to the formulation given.

We conclude this section with some remarks on symmetry under time-reversal. The following calculation shows that $J_1 = K_0^*$ and hence that the cut locus is symmetric. Since V is Hamiltonian, its flow preserves the symplectic form β . See for example [20]. For $\xi \in T^*M$, write $\tilde{\xi}$ for the corresponding vertical vector in TT^*M and write ψ_t^* for the action of ψ_t on TT^*M . Then

$$\langle J_1 \xi_0, \xi_1 \rangle = \langle \pi^* \psi_1^* \tilde{\xi}_0, \xi_1 \rangle = \beta(\psi_1^* \tilde{\xi}_0, \tilde{\xi}_1) = \beta(\tilde{\xi}_0, \psi_{-1}^* \tilde{\xi}_1) = -\langle \xi_0, \pi^* \psi_{-1}^* \tilde{\xi}_1 \rangle = \langle \xi_0, K_0 \xi_1 \rangle.$$

When \mathcal{L} has the form (8), for compactly supported C^∞ functions f, g on M , we have

$$\int_M f \mathcal{L} g \, d\nu = \int_M g \hat{\mathcal{L}} f \, d\nu.$$

where

$$\hat{\mathcal{L}} f = \frac{1}{2} \operatorname{div}(a \nabla f) - a(\beta, \nabla f).$$

and the associated bridge measure satisfies $\hat{\mu}_\varepsilon^{y,x} = \mu_\varepsilon^{x,y} \circ \wedge^{-1}$, where $\wedge : \Omega^{x,y} \rightarrow \Omega^{y,x}$ is the time-reversal map, given by $\hat{\omega}_t = \omega_{1-t}$. Hence, in this case, the whole set-up is invariant under time-reversal.

3 Brownian bridge on a Riemannian manifold

Before proving Theorem 1.3, we discuss its consequences in the case where a is everywhere positive-definite. Let M be a connected C^∞ Riemannian manifold. Write Δ for the Laplace–Beltrami operator on M and set

$$\mathcal{L} = \frac{1}{2}\Delta + \bar{X}_0$$

where \bar{X}_0 is a C^∞ vector field on M . Then \mathcal{L} has the form (4) and its diffusivity a is the inverse of the metric tensor. We obtain in this way all operators \mathcal{L} of the form (4) with a everywhere positive-definite. In this context, it is well known that there exist $m \in \mathbb{N}$ and C^∞ vector fields X_1, \dots, X_m on M such that (1) and (2) hold, with

$$\text{span}\{Y(x) : Y \in \mathcal{A}(X_1, \dots, X_m)\} = \text{span}\{X_1(x), \dots, X_m(x)\} = T_x M \quad \text{for all } x \in M.$$

We now review a number of equivalent characterizations of the cut locus for Riemannian manifolds, and we give some further descriptions of the limit Gaussian measure μ_γ using objects of Riemannian geometry. In the case $\bar{X}_0 = 0$, the measure $\mu_\varepsilon^{x,y}$ is the law of the Riemannian Brownian bridge from x to y of speed $\sqrt{\varepsilon}$.

Fix $x, y \in M$ and assume, as above, that there is a unique strongly minimal path $\gamma \in H^{x,y}$. It is well known that in this context γ is always the projection of a bicharacteristic. Write ∇ for the Levi–Civita covariant derivative and R for the Riemann curvature tensor. Define a linear map $R_t : T_{\gamma_t} M \rightarrow T_{\gamma_t} M$, symmetric with respect to the metric, by $R_t = R(\cdot, \dot{\gamma}_t)\dot{\gamma}_t$. It is straightforward to see that the set $T_\gamma H^{x,y}$, defined more generally in Section 7, is here given by the set of all absolutely continuous paths v in $T_\gamma \Omega^{x,y}$ such that

$$\int_0^1 |\nabla v_t|^2 dt < \infty.$$

Moreover, we can define, for $\eta \in \mathbb{R}$ sufficiently small, a path $\gamma^\eta \in H^{x,y}$ by $\gamma_t^\eta = \exp_{\gamma_t}(\eta v_t)$. Then the map $\eta \mapsto I(\gamma^\eta)$ is twice differentiable near 0 with

$$Q(v) = \left. \frac{\partial^2}{\partial \eta^2} \right|_{\eta=0} I(\gamma^\eta) = \int_0^1 |\nabla v_t|^2 dt - \int_0^1 \langle v_t, R_t v_t \rangle dt. \quad (18)$$

See, for example [16]. By a standard calculation, the processes $(J_t)_{t \in [0,1]}$ and $(K_t)_{t \in [0,1]}$, defined above, are Jacobi fields along γ and satisfy the differential equations

$$\nabla^2 J_t + R_t J_t = 0, \quad J_0 = 0, \quad \nabla J_0 = a(x)$$

and

$$\nabla^2 K_t + R_t K_t = 0, \quad K_1 = 0, \quad \nabla K_1 = -a(y).$$

Let $(b_t)_{t \in [0,1]}$ be a Brownian motion in $T_x M$, starting from 0. Set $z_t = b_t - tb_1$. Then $(z_t)_{t \in [0,1]}$ is a Brownian bridge in $T_x M$ from 0 to 0 in time 1. Let $(\tau_t)_{t \in [0,1]}$ denote parallel translation along γ , thus $\tau_t \in T_{\gamma_t} M \otimes T_x^* M$ and $\tau_0 = \text{id}$, $\nabla \tau_t = 0$. Let $\bar{\mu}$ denote the law of $(\tau_t z_t)_{t \in [0,1]}$ on $T_\gamma \Omega^{x,y}$. The following result is well known. See [16] and, for (vi), [7, Theorem 4.17]. Although it is framed geometrically, it is in fact only a result about quadratic equations in matrices and Gaussian processes, as one can see by choosing a chart along γ such that $a(\gamma_t) = \tau_t = \text{id}$ for all t .

Proposition 3.1. *The following are equivalent*

- (i) J_t is invertible for all $t \in (0, 1]$,
- (ii) K_t is invertible for all $t \in [0, 1)$,
- (iii) there exists a C^1 path $(A_t)_{t \in [0,1]}$ along γ , with $A_t \in T_{\gamma_t} M \otimes T_{\gamma_t}^* M$, solving the Riccati equation

$$\nabla A_t + A_t^2 + R_t = 0, \quad (1-t)A_t \rightarrow -\text{id} \quad \text{as } t \rightarrow 1,$$

- (iv) there exists a C^1 path $(B_t)_{t \in (0,1]}$ along γ , with $B_t \in T_{\gamma_t} M \otimes T_{\gamma_t}^* M$, solving the Riccati equation

$$\nabla B_t + B_t^2 + R_t = 0, \quad tB_t \rightarrow \text{id} \quad \text{as } t \rightarrow 0,$$

- (v) Q is positive-definite on $T_\gamma H^{x,y}$,

- (vi) we have

$$\int_{T_\gamma \Omega^{x,y}} \exp \left\{ \frac{1}{2} \int_0^1 \langle v_t, R_t v_t \rangle dt \right\} \bar{\mu}(dv) < \infty.$$

Moreover, under these conditions, we have

$$J_1 = K_0^*, \quad \nabla K_t = A_t K_t, \quad \nabla J_t = B_t J_t$$

and for $t < 1$

$$J_t = K_t \int_0^t K_s^{-1} a(\gamma_s) (K_s^{-1})^* ds K_0^*.$$

It is standard that the condition that γ is minimal implies already that Q is non-negative and that J_t is invertible for all $t \in (0, 1)$. In this context then, condition (i) is equivalent to the condition that x and y are non-conjugate along γ , that is, there is no non-trivial vector field $(v_t)_{t \in [0,1]}$ along γ vanishing at the endpoints and such that $\nabla^2 v_t + R_t v_t = 0$ for all t . In Section 7, we will generalize the equivalence of (i), (ii) and (v) to the sub-Riemannian case.

The following result gives three further characterizations for the Gaussian measure μ_γ , which by Theorem 1.3 describes the small-time fluctuations of the Brownian bridge in M . As in the preceding result, by choice of a suitable chart along γ , we can reduce to the case where $M = \mathbb{R}^d$ with $a(\gamma_t) = \tau_t = \text{id}$ for all t . The result is then of a standard type for Gaussian processes. See for example [24].

Theorem 3.2. *Let M be a connected Riemannian manifold and let $x, y \in M$. Suppose that there is a unique minimal path $\gamma \in H^{x,y}$ and that (x, y) is non-conjugate along γ . Then there exists a zero-mean Gaussian probability measure μ_γ on $T_\gamma\Omega^{x,y}$ with the following properties:*

(i) *for $s \leq t$, we have*

$$\int_{T_\gamma\Omega^{x,y}} v_s \otimes v_t \mu_\gamma(dv) = J_s J_1^{-1} K_t^*,$$

(ii) *for all continuous linear functionals ϕ on $T_\gamma\Omega^{x,y}$, we have*

$$\int_{T_\gamma\Omega^{x,y}} \phi(v)^2 \mu_\gamma(dv) = Q(\tilde{\phi})$$

where $\tilde{\phi} \in T_\gamma H^{x,y}$ is determined by $\phi(v) = Q(\tilde{\phi}, v)$ for all $v \in T_\gamma H^{x,y}$,

(iii) *under μ_γ , the coordinate process v on $T_\gamma\Omega^{x,y}$ satisfies a covariant linear stochastic differential equation over γ of the form*

$$\nabla v_t = \tau_t db_t + A_t v_t dt, \quad v_0 = 0,$$

(iv) *μ_γ is absolutely continuous with respect to $\bar{\mu}$, with Radon–Nikodym derivative*

$$\frac{d\mu_\gamma}{d\bar{\mu}}(v) \propto \exp \left\{ \frac{1}{2} \int_0^1 \langle v_t, R_t v_t \rangle dt \right\}. \quad (19)$$

Moreover, any one of these properties characterizes μ_γ uniquely.

The equivalence of (i) and (ii) will be established in a more general context in Section 8. Here is an illustrative calculation, taking advantage of the reduction to $\tau_t = \text{id}$ mentioned above. Suppose that μ satisfies (iii). Set $w_t = K_t^{-1} v_t$, then $dw_t = K_t^{-1} db_t$, so

$$w_t = \int_0^t K_s^{-1} db_s.$$

Hence, for $0 \leq s \leq t \leq 1$,

$$\int_{T_\gamma\Omega^{x,y}} v_s \otimes v_t \mu(dv) = K_s \left(\int_0^s K_r^{-1} (K_r^{-1})^* dr \right) K_t^* = J_s J_1^{-1} K_t^*.$$

Hence μ satisfies (i).

We examine now how our result specializes in some simple cases. When $M = \mathbb{R}^d$ with Euclidean metric, the analysis is trivial, because $\tilde{\mu}_\varepsilon^{x,y} = \mu_\gamma$ for all $\varepsilon > 0$. Then

$$J_s = sI, \quad K_t = (1-t)I, \quad R_t = 0, \quad A_t = -I/(1-t)$$

so the alternatives in Theorem 3.2 are simply some of the standard descriptions of the Brownian bridge in \mathbb{R}^d .

In the case where M is a sphere or hyperbolic space, we can rewrite (19) in the form

$$\frac{d\mu_\gamma}{d\bar{\mu}}(y) \propto \exp \left\{ \frac{Kd(x,y)^2}{2} \int_0^1 |y_t|^2 dt \right\},$$

where K , the scalar curvature, is 1 for the sphere and -1 for hyperbolic space. Thus, on a sphere, the variance of the fluctuations is larger than in \mathbb{R}^d , whereas, in hyperbolic space it is less. This does not contradict the tendency of Brownian paths to separate quickly in hyperbolic space because we are conditioning on the endpoint. Thus we tend to see those paths which have never deviated far from the geodesic.

4 Laplace's method on Wiener space

In this section we prove Theorem 1.3 in the case where conditions (5) and (6) hold, following closely the method used by Ben Arous [6] to study the heat kernel in this case. It will suffice to consider the case where $\theta(z) = z$ for all $z \in \mathbb{R}^d$. The vector fields X_0, X_1, \dots, X_m provide both a means to construct a diffusion process $(x_t^\varepsilon)_{t \geq 0}$ with generator $\varepsilon \mathcal{L}$ starting from x and a parametrization of the set H^x of finite energy paths starting from x . Let $(B_t)_{t \in [0,1]}$ be a Brownian motion in \mathbb{R}^m , which we assume is realized as the coordinate process on $\Omega^0(\mathbb{R}^m) = \{w \in C([0,1], \mathbb{R}^m) : w_0 = 0\}$ under Wiener measure \mathbb{P} . Define a vector field \tilde{X}_0 on \mathbb{R}^d by

$$\tilde{X}_0^i(x) = X_0^i(x) + \frac{1}{2} \sum_{\ell=1}^m \langle \nabla X_\ell^i(x), X_\ell(x) \rangle.$$

Consider the Itô stochastic differential equation in \mathbb{R}^d

$$dx_t^\varepsilon = \sqrt{\varepsilon} \sum_{\ell=1}^m X_\ell(x_t^\varepsilon) dB_t^\ell + \varepsilon \tilde{X}_0(x_t^\varepsilon) dt, \quad x_0^\varepsilon = x.$$

This has a unique strong solution $(x_t^\varepsilon)_{t \in [0,1]}$, whose law on $\Omega = C([0,1], \mathbb{R}^d)$ is μ_ε^x . There is no explosion. Write $H^0(\mathbb{R}^m)$ for the set of Cameron–Martin paths h in \mathbb{R}^m , that is to say, the

set of absolutely continuous functions $h : [0, 1] \rightarrow \mathbb{R}^m$, starting from 0, such that

$$\int_0^1 |\dot{h}_t|^2 dt < \infty.$$

Given $h \in H^0(\mathbb{R}^m)$, consider the differential equation in \mathbb{R}^d

$$d\phi_t = \sum_{\ell=1}^m X_\ell(\phi_t) dh_t^\ell, \quad \phi_0 = x. \quad (20)$$

There is a unique solution $\phi(x, h) = (\phi_t(x, h))_{t \in [0, 1]}$ in Ω . As we will show in Proposition 7.1, in fact $\phi(x, h) \in H^x$ and $\phi(x, \cdot)$ maps $H^0(\mathbb{R}^m)$ onto H^x . Recall that γ is the minimizing path in $H^{x, y}$ and γ is the projection of a bicharacteristic λ . Define $h \in H^0(\mathbb{R}^m)$ by $\dot{h}_t^\ell = \langle \lambda_t, X_\ell(\gamma_t) \rangle$. Then $\gamma = \phi(x, h)$ and $I(\gamma) = \int_0^1 |\dot{h}_t|^2 dt$. We reserve the notation h for this minimizing control path from now on. For $s \in [0, 1]$, we will write $(\phi_{ts}(x, h))_{t \in [s, 1]}$ for the solution to (20) starting from x at time s . We denote the derivative in x by $\phi_{ts}^*(x, h)$ and set $u_t = \phi_t^*(x, h)$. Then u_t is invertible for all t and

$$du_t = \sum_{\ell=1}^m \nabla X_\ell(\gamma_t) u_t dh_t^\ell, \quad u_0 = I.$$

Moreover, we have $\phi_{ts}^*(\gamma_s, h) = u_t u_s^{-1}$. We can define a continuous linear map $v : \Omega^0(\mathbb{R}^m) \rightarrow T_\gamma \Omega^{x, y}$ by⁷

$$v_t(w) = \sum_{\ell=1}^m \int_0^t \phi_{ts}^*(\gamma_s, h) X_\ell(\gamma_s) dw_s^\ell = \sum_{\ell=1}^m u_t \int_0^t u_s^{-1} X_\ell(\gamma_s) dw_s^\ell \quad (21)$$

where the integral is understood by a formal integration by parts. Set $Y_t = v_t(B)$ and note that Y_1 is a zero-mean Gaussian random variable in \mathbb{R}^d having covariance matrix $\bar{C}_1 = u_1 C_1 u_1^*$, where

$$C_1 = \sum_{\ell=1}^m \int_0^1 (u_t^{-1} X_\ell(\gamma_t)) \otimes (u_t^{-1} X_\ell(\gamma_t)) dt.$$

In [7], Bismut called C_1 the deterministic Malliavin covariance matrix. Condition (16) implies that C_1 is invertible. This follows in particular from (61). Hence Y_1 has a density function with respect to Lebesgue measure on \mathbb{R}^d , given by

$$\bar{p}(z) = (2\pi)^{-d/2} (\det \bar{C}_1)^{-1/2} \exp\{-\langle z, \bar{C}_1^{-1} z \rangle / 2\}.$$

⁷Since $M = \mathbb{R}^d$ in the present discussion, $T_\gamma \Omega^{x, y}$ can be naturally identified with a subset of Ω , but we will keep the distinction anyway.

Set⁸

$$Y_1^{(2)} = \sum_{\ell=1}^m \int_0^1 \phi_{1t}^*(\gamma_t, h) \{ \nabla^2 X_\ell(\gamma_t)(Y_t, Y_t) dh_t^\ell + 2 \nabla X_\ell(\gamma_t) Y_t dB_t^\ell \}, \quad S = \langle \lambda_1, Y_1^{(2)} \rangle.$$

where the integral $Y_t dB_t$ is understood in the sense of Itô. Define a linear map $\tau : \mathbb{R}^d \rightarrow H^0(\mathbb{R}^m)$ by

$$\dot{\tau}_t^\ell(z) = \langle C_1^{-1} u_1^{-1} z, u_t^{-1} X_\ell(\gamma_t) \rangle.$$

Set $K = \{k \in H^0(\mathbb{R}^m) : v_1(k) = 0\}$. From (21) we see that τ maps \mathbb{R}^d onto the orthogonal complement K^\perp of K in $H^0(\mathbb{R}^m)$. Moreover, $v_1(\tau(z)) = z$ and the restriction of the map $\tau \circ v_1$ to $H^0(\mathbb{R}^m)$ is the orthogonal projection $H^0(\mathbb{R}^m) \rightarrow K^\perp$. Set

$$W_t = B_t - W'_t, \quad W'_t = \tau_t(v_1(B)). \quad (22)$$

Then $(W_t)_{t \in [0,1]}$ and $(W'_t)_{t \in [0,1]}$ are independent continuous Gaussian processes, and $v_1(W) = 0$ and $W' \in K^\perp$ almost surely. For $z \in \mathbb{R}^d$, set $W_t(z) = W_t + \tau_t(z)$ and $Y_t(z) = v_t(W(z))$ and $S(z) = \langle \lambda_1, Y_1^{(2)}(z) \rangle$, where

$$Y_1^{(2)}(z) = \sum_{\ell=1}^m \int_0^1 \phi_{1t}^*(\gamma_t, h) \{ \nabla^2 X_\ell(\gamma_t)(Y_t(z), Y_t(z)) dh_t^\ell + 2 \nabla X_\ell(\gamma_t) Y_t(z) dW_t^\ell(z) \}. \quad (23)$$

We interpret the integral with respect to $W(z)$ by writing

$$Y_t(z) dW_t(z) = v_t(B - W' + \tau(z)) d(B_t - W'_t + \tau_t(z))$$

and expanding. The term $v_t(B) dB_t$ is considered as an Itô integral, while the remaining terms can be considered as integrals with respect to Lebesgue measure, sometimes after a formal integration by parts. It is straightforward to see, using the same version of the Itô integral for all z , that the family of random variables $(Y_1^{(2)}(z) : z \in \mathbb{R}^d)$ is continuous in z . Also, $Y(z)$ and $Y_1^{(2)}(z)$ are independent of $v_1(B)$ for all z . Note that $W(v_1(B)) = B$ and $Y(v_1(B)) = Y$, and that $S(v_1(B)) = S$ almost surely. Note also that $Y_1^{(2)}(z)$ belongs to the sum of the zeroth, first and second Wiener chaoses in $L^2(\Omega^0(\mathbb{R}^m), \mathbb{P})$ for all z . By Theorem 8.5, we have $\mathbb{E}(e^{pS(z)/2}) < \infty$ for all $z \in \mathbb{R}^d$ for some $p > 1$.

Consider the function

$$E(z) = d(x, z)^2/2 = \inf_{\omega \in H^{x,z}} I(\omega)/2.$$

⁸Formally, we have

$$Y_t = \frac{\partial}{\partial h} \phi_t(x, h)(B), \quad Y_1^{(2)} = \frac{\partial^2}{\partial h^2} \phi_1(x, h)(B, B)$$

which may be seen by differentiating (20) twice in h in the direction B and solving the resulting equations by variation of constants.

As Bismut [7, Theorem 1.26] showed, E is C^∞ in a neighbourhood N of y , with $E'(y) = \lambda_1$. Following Ben Arous [6, Lemma 3.8], there then exists a function $F \in C_b^\infty(\mathbb{R}^d)$ such that the map $F + E$ has a unique and non-degenerate minimum at y , with minimum value 0. To see this, choose a neighbourhood N_0 of y and a C^∞ function χ of compact support such that $1_{N_0} \leq \chi \leq 1_N$. Fix a constant $\alpha > 0$ and consider the function

$$F(z) = \chi(z) (\alpha|y - z|^2 - E(y) - E'(y)(y - z)) + (1 - \chi(z)).$$

Then $F \in C_b^\infty(\mathbb{R}^d)$ and $F(y) + E(y) = 0$ and $F'(y) + E'(y) = 0$. Moreover, by choosing α sufficiently large, we can ensure that $F''(y) + E''(y)$ is positive-definite and $F(z) + E(z) > 0$ for all $z \neq y$, so $F + E$ has a non-degenerate minimum at y , which is also its global minimum. We fix a choice of a function F with the given properties for the rest of the analysis.

Set $\gamma^0 = \gamma$ and define $(\gamma_t^\varepsilon)_{t \in [0,1]}$ for $\varepsilon > 0$ as the strong solution of the stochastic differential equation

$$d\gamma_t^\varepsilon = \sum_{\ell=1}^m X_\ell(\gamma_t^\varepsilon) dh_t^\ell + \sqrt{\varepsilon} \sum_{\ell=1}^m X_\ell(\gamma_t^\varepsilon) dB_t^\ell + \varepsilon \tilde{X}_0(\gamma_t^\varepsilon) dt, \quad \gamma_0^\varepsilon = x. \quad (24)$$

By standard results on stochastic differential equations, for all $t \in [0, 1]$ and all $p \in [1, \infty)$, the map $\varepsilon \mapsto \gamma_t^\varepsilon : [0, \infty) \rightarrow L^p(\mathbb{P})$ is continuous. Furthermore, we can and do choose versions so that, almost surely, the map $\sigma \mapsto \gamma_t^{\sigma^2} : [0, \infty) \rightarrow \mathbb{R}^d$ is C^∞ . Moreover, the first and second derivatives at $\sigma = 0$ then satisfy

$$\left. \frac{\partial}{\partial \sigma} \right|_{\sigma=0} \gamma_t^{\sigma^2} = Y_t, \quad \left(\left. \frac{\partial}{\partial \sigma} \right|_{\sigma=0} \right)^2 \gamma_1^{\sigma^2} = Y_1^{(2)} + Z_1, \quad Z_1 = \int_0^1 \phi_{1t}^*(\gamma_t, h) \tilde{X}_0(\gamma_t) dt.$$

Now the map $f(\sigma) = F(\gamma_1^{\sigma^2})$ is C^∞ on $[0, \infty)$ and $F(y) = -d(x, y)^2/2$ and $F'(y) = -\lambda_1$, so

$$f(0) = F(y) = -\frac{1}{2} \int_0^1 |\dot{h}_t|^2 dt, \quad f'(0) = F'(y) Y_1 = -\sum_{\ell=1}^m \int_0^1 \dot{h}_t^\ell dB_t^\ell$$

and

$$f''(0) = F'(y)(Y_1^{(2)} + Z_1) + F''(y)(Y_1, Y_1).$$

Set

$$R(\varepsilon) = \int_0^1 (1 - \theta) f''(\theta \sqrt{\varepsilon}) d\theta.$$

Then, by Taylor's theorem,

$$F(\gamma_1^\varepsilon) = f(\sqrt{\varepsilon}) = f(0) + \sqrt{\varepsilon} f'(0) + \varepsilon R(\varepsilon) = -\frac{1}{2} \int_0^1 |\dot{h}_t|^2 dt - \sqrt{\varepsilon} \sum_{\ell=1}^m \int_0^1 \dot{h}_t^\ell dB_t^\ell + \varepsilon R(\varepsilon).$$

Set

$$\tilde{x}_t^\varepsilon = \frac{x_t^\varepsilon - \gamma_t}{\sqrt{\varepsilon}}, \quad \tilde{\gamma}_t^\varepsilon = \frac{\gamma_t^\varepsilon - \gamma_t}{\sqrt{\varepsilon}}$$

and note that, for all $t \in [0, 1]$, we have $\tilde{\gamma}_t^\varepsilon \rightarrow Y_t$ as $\varepsilon \rightarrow 0$ almost surely. For $\varepsilon > 0$, consider the new probability measure \mathbb{P}^ε on $\Omega^0(\mathbb{R}^m)$ given by $d\mathbb{P}^\varepsilon/d\mathbb{P} = \rho_1^\varepsilon$, where

$$\rho_1^\varepsilon = \exp \left\{ -\frac{1}{\sqrt{\varepsilon}} \sum_{\ell=1}^m \int_0^1 \dot{h}_t^\ell dB_t^\ell - \frac{1}{2\varepsilon} \int_0^1 |\dot{h}_t|^2 dt \right\} = \exp \left\{ \frac{F(\gamma_1^\varepsilon)}{\varepsilon} - R(\varepsilon) \right\}.$$

Note that (24) can be rewritten in the form

$$d\gamma_t^\varepsilon = \sqrt{\varepsilon} \sum_{\ell=1}^m X_\ell(\gamma_t^\varepsilon) dB_t^{\varepsilon, \ell} + \varepsilon \tilde{X}_0(\gamma_t^\varepsilon) dt, \quad \gamma_0^\varepsilon = x$$

where $B_t^{\varepsilon, \ell} = B_t^\ell + h_t^\ell/\sqrt{\varepsilon}$. By the Cameron-Martin formula, under \mathbb{P}^ε , the process $(B_t^\varepsilon)_{t \in [0, 1]}$ is a Brownian motion, so γ^ε has law μ_ε^x .

At this point we modify the argument of Ben Arous by introducing a smooth cylindrical function G on Ω which serves to keep track of the paths of the diffusion bridge. Fix $t_1, \dots, t_k \in (0, 1)$ with $t_1 < \dots < t_k$ and a C^∞ function g on $(\mathbb{R}^d)^k$ of polynomial growth. Set $G(\omega) = g(\omega_{t_1}, \dots, \omega_{t_k})$. Define for $z \in \mathbb{R}^d$ and $\varepsilon > 0$

$$G_\varepsilon(z) = \varepsilon^{d/2} p(\varepsilon, x, y + \sqrt{\varepsilon}z) e^{-F(y + \sqrt{\varepsilon}z)/\varepsilon} \int_{T_\gamma \Omega^{x, y}} G(\omega) \tilde{\mu}_\varepsilon^{x, y + \sqrt{\varepsilon}z}(d\omega),$$

$$G_0(z) = \bar{p}(z) e^{\langle \lambda_1, Z_1 \rangle / 2 - F''(y)(z, z)/2} \mathbb{E}(G(Y(z)) e^{S(z)/2}).$$

Then G_ε and G_0 are continuous integrable functions on \mathbb{R}^d . Consider the Fourier transform

$$\hat{G}_\varepsilon(\xi) = \int_{\mathbb{R}^d} G_\varepsilon(z) e^{i\langle \xi, z \rangle} dz.$$

Then

$$\begin{aligned} \hat{G}_\varepsilon(\xi) &= \int_{\mathbb{R}^d} \int_{\Omega} p(\varepsilon, x, y') e^{-F(y')/\varepsilon} G\left(\frac{\omega - \gamma}{\sqrt{\varepsilon}}\right) \mu_\varepsilon^{x, y'}(d\omega) e^{i\langle \xi, (y' - y)/\sqrt{\varepsilon} \rangle} dy' \\ &= \int_{\Omega} e^{-F(\omega_1)/\varepsilon} G\left(\frac{\omega - \gamma}{\sqrt{\varepsilon}}\right) e^{i\langle \xi, (\omega_1 - \gamma_1)/\sqrt{\varepsilon} \rangle} \mu_\varepsilon^x(d\omega) \\ &= \mathbb{E}(G(\tilde{x}^\varepsilon) \exp\{i\langle \xi, \tilde{x}_1^\varepsilon \rangle - F(x_1^\varepsilon)/\varepsilon\}) \\ &= \mathbb{E}(G(\tilde{\gamma}^\varepsilon) \exp\{i\langle \xi, \tilde{\gamma}_1^\varepsilon \rangle - F(\gamma_1^\varepsilon)/\varepsilon\} \rho_1^\varepsilon) \\ &= \mathbb{E}(G(\tilde{\gamma}^\varepsilon) \exp\{i\langle \xi, \tilde{\gamma}_1^\varepsilon \rangle - R(\varepsilon)\}) \end{aligned} \tag{25}$$

and

$$\begin{aligned}
\hat{G}_0(\xi) &= \int_{\mathbb{R}^d} \bar{p}(z) e^{\langle \lambda_1, Z_1 \rangle / 2 - F''(y)(z, z) / 2} \mathbb{E}(G(Y(z)) e^{S(z) / 2}) e^{i \langle \xi, z \rangle} dz \\
&= \mathbb{E}(G(Y) \exp\{i \langle \xi, Y_1 \rangle + (S + \langle \lambda_1, Z_1 \rangle - F''(y)(Y_1, Y_1)) / 2\}) \\
&= \mathbb{E}(G(Y) \exp\{i \langle \xi, Y_1 \rangle - f''(0) / 2\}).
\end{aligned}$$

Consider the limit $\varepsilon \rightarrow 0$. We have $\tilde{\gamma}_t^\varepsilon \rightarrow Y_t$ almost surely and in L^p for all $p < \infty$, for all $t \in [0, 1]$. Now $R(\varepsilon) \rightarrow f''(0)/2$ almost surely as $\varepsilon \rightarrow 0$ and we have the following key estimate proved by Ben Arous [6, Lemma 3.25]: there exists $p > 1$ such that

$$\limsup_{\varepsilon \rightarrow 0} \mathbb{E}(e^{-pR(\varepsilon)}) < \infty. \quad (26)$$

Hence $\hat{G}_\varepsilon(\xi) \rightarrow \hat{G}_0(\xi)$ for all $\xi \in \mathbb{R}^d$. We will prove the following key lemma at the end of this section.

Lemma 4.1. *There exists a constant $C(G) < \infty$ such that, for all $\varepsilon \in (0, 1]$ and all $\xi \in \mathbb{R}^d$, we have*

$$|\hat{G}_\varepsilon(\xi)| \leq C(G) / (1 + |\xi|^{d+1}). \quad (27)$$

Moreover, there exists $C < \infty$ such that, uniformly in $s, t \in [0, 1]$, in the case where $G(\omega) = |\omega_s - \omega_t|^4$, for all $\varepsilon \in (0, 1]$ and all $\xi \in \mathbb{R}^d$, we have

$$|\hat{G}_\varepsilon(\xi)| \leq C |s - t|^2 / (1 + |\xi|^{d+1}). \quad (28)$$

The lemma allows us to use dominated convergence in the Fourier inversion formula to deduce that

$$G_\varepsilon(0) = (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{G}_\varepsilon(\xi) d\xi \rightarrow (2\pi)^{-d} \int_{\mathbb{R}^d} \hat{G}_0(\xi) d\xi = G_0(0).$$

that is to say

$$\varepsilon^{d/2} p(\varepsilon, x, y) e^{d(x, y)^2 / (2\varepsilon)} \int_{T_\gamma \Omega^{x, y}} G(\omega) \tilde{\mu}_\varepsilon^{x, y}(d\omega) \rightarrow (2\pi)^{-d/2} (\det \bar{C}_1)^{-1/2} e^{\langle \lambda_1, Z_1 \rangle / 2} \mathbb{E}(G(Y(0)) e^{S(0) / 2}) \quad (29)$$

where we used the fact that $\bar{p}(0) = (2\pi)^{-d/2} (\det \bar{C}_1)^{-1/2}$. On taking $g = 1$, we recover the heat kernel asymptotics shown by Ben Arous

$$\varepsilon^{d/2} p(\varepsilon, x, y) e^{d(x, y)^2 / (2\varepsilon)} \rightarrow (2\pi)^{-d/2} (\det \bar{C}_1)^{-1/2} e^{\langle \lambda_1, Z_1 \rangle / 2} \mathbb{E}(e^{S(0) / 2}). \quad (30)$$

By Theorem 8.5, μ_γ is the law of $Y(0)$ on $T_\gamma \Omega^{x, y}$ under the probability measure $\tilde{\mathbb{P}}$, where $d\tilde{\mathbb{P}}/d\mathbb{P} \propto e^{S(0)/2}$. Hence, on dividing (29) by (30), we obtain

$$\int_{T_\gamma \Omega^{x, y}} G(\omega) \tilde{\mu}_\varepsilon^{x, y}(d\omega) \rightarrow \int_{T_\gamma \Omega^{x, y}} G(\omega) \mu_\gamma(d\omega)$$

from which we deduce that the finite-dimensional distributions of $\tilde{\mu}_\varepsilon^{x,y}$ converge weakly to those of μ_γ . Finally, on taking $G(\omega) = |\omega_s - \omega_t|^4$ and using the estimate (28), we find by Fourier inversion that

$$\varepsilon^{d/2} p(\varepsilon, x, y) e^{d(x,y)^2/(2\varepsilon)} \int_{T_\gamma \Omega^{x,y}} |\omega_s - \omega_t|^4 \tilde{\mu}_\varepsilon^{x,y}(d\omega) = G_\varepsilon(0) \leq C|s - t|^2.$$

Since the right-hand side in (30) is positive, we deduce that for some $C < \infty$ we have, for all $s, t \in [0, 1]$,

$$\sup_{\varepsilon \in (0,1]} \int_{T_\gamma \Omega^{x,y}} |\omega_s - \omega_t|^4 \tilde{\mu}_\varepsilon^{x,y}(d\omega) \leq C|s - t|^2.$$

Hence, by standard arguments, the family of laws $(\tilde{\mu}_\varepsilon^{x,y} : \varepsilon \in (0, 1])$ is tight on $T_\gamma \Omega^{x,y}$ and so $\tilde{\mu}_\varepsilon^{x,y} \rightarrow \mu_\gamma$ weakly as $\varepsilon \rightarrow 0$. In particular, we also have $\mu_\varepsilon^{x,y} \rightarrow \delta_\gamma$, so the proof is complete.

Proof of Lemma 4.1. The idea is to integrate by parts in (25), $d + 1$ times, using Malliavin calculus. We will vary the argument of [6] in three respects. First, we will present the argument using Bismut's integration by parts formula [8, 25] for solutions of stochastic differential equations. This is more direct and may be followed in detail without knowledge of the general apparatus of Malliavin calculus. Second, we will use the corrected argument [21] for the uniform non-degeneracy of the Malliavin covariance matrix. Third, we will include the simple modifications needed to go beyond the case $G = 1$ which is covered in [6]. Within the proof, we will use a few notations which conflict with usage elsewhere.

Fix $\sigma = \sqrt{\varepsilon} > 0$. Define processes $(x_t)_{t \in [0,1]}$ in \mathbb{R}^d and $(u_t)_{t \in [0,1]}$, $(v_t)_{t \in [0,1]}$ in $\mathbb{R}^d \otimes (\mathbb{R}^d)^*$ as the strong solutions of the following system of stochastic differential equations

$$\begin{aligned} dx_t &= \sum_{\ell=1}^m X_\ell(x_t) dh_t^\ell + \sum_{\ell=1}^m \sigma X_\ell(x_t) dB_t^\ell + \sigma^2 \tilde{X}_0(x_t) dt, \quad x_0 = x \\ du_t &= \sum_{\ell=1}^m \nabla X_\ell(x_t) u_t dh_t^\ell + \sum_{\ell=1}^m \sigma \nabla X_\ell(x_t) u_t dB_t^\ell + \sigma^2 \nabla \tilde{X}_0(x_t) u_t dt, \quad u_0 = I \\ dv_t &= - \sum_{\ell=1}^m v_t \nabla X_\ell(x_t) dh_t^\ell - \sum_{\ell=1}^m \sigma v_t \nabla X_\ell(x_t) dB_t^\ell - \sigma^2 v_t \left\{ \nabla \tilde{X}_0 - \sum_{\ell=1}^m (\nabla X_\ell)^2 \right\} (x_t) dt, \quad v_0 = I. \end{aligned} \tag{31}$$

Then $x_t = \gamma_t^{\sigma^2}$ and, by Itô's formula, $v_t = u_t^{-1}$ for all t . It is well known that the random variables $\sup_{t \in [0,1]} |x_t|$, $\sup_{t \in [0,1]} |u_t|$ and $\sup_{t \in [0,1]} |v_t|$ have moments of all orders, which are bounded uniformly in $\sigma \in [0, 1]$.

Consider the following stochastic differential equation in \mathbb{R}^N with C^∞ coefficients

$$dz_t = \sum_{\ell=1}^m W_\ell(z_t) dh_t^\ell + \sum_{\ell=1}^m Z_\ell(z_t) dB_t^\ell + Z_0(z_t) dt, \quad z_0 = z. \tag{32}$$

We assume that the coefficients have a graded Lipschitz structure. By this we mean that the coefficients and all their derivatives satisfy polynomial growth bounds, and that there exist $k \in \mathbb{N}$ and $N_1, \dots, N_k \in \mathbb{N}$ and a decomposition $z_t = (z_t^1, \dots, z_t^k)$ such that $N_1 + \dots + N_k = N$, and, for $j = 1, \dots, k$, the component z_t^j takes values in \mathbb{R}^{N_j} and, for all ℓ , the corresponding components W_ℓ^j and Z_ℓ^j of the coefficients depend only on (z^1, \dots, z^j) , with $\partial W_\ell^j / \partial z^j$ and $\partial Z_\ell^j / \partial z^j$ uniformly bounded. We will allow the coefficients in (33) to depend measurably on σ and t , without making this explicit in the notation, while assuming that the bounds in their graded Lipschitz structure hold uniformly in $\sigma \in [0, 1]$ and $t \in [0, 1]$. This is sufficient to ensure the existence and uniqueness of a strong solution $(z_t)_{t \in [0, 1]}$ to (33) such that $\sup_{t \in [0, 1]} |z_t|$ has moments of all orders, bounded uniformly in $\sigma \in [0, 1]$. See [25] for more details.

Fix $i \in \{1, \dots, d\}$ and consider for $\eta \in \mathbb{R}$ a perturbed process $(B_t^\eta)_{t \in [0, 1]}$ in \mathbb{R}^m given by

$$dB_t^{\eta, \ell} = dB_t^\ell + \eta(v_t X_\ell(x_t))^i dt, \quad B_0^\eta = 0.$$

Write $(z_t^\eta)_{t \in [0, 1]}$ for the strong solution of the stochastic differential equation which is obtained when we replace $(B_t)_{t \in [0, 1]}$ in (33) by this perturbed process. We can and do choose a version of the family of processes $(z_t^\eta)_{t \in [0, 1]}$ which is almost surely C^∞ in η . We associate to $(z_t)_{t \in [0, 1]}$ the derived process $(z'_t)_{t \in [0, 1]} = ((z'_t)^1, \dots, (z'_t)^d)_{t \in [0, 1]}$ in $\mathbb{R}^N \otimes \mathbb{R}^d$ given by

$$(z'_t)^i = (\partial / \partial \eta)|_{\eta=0} z_t^\eta. \quad (34)$$

Then $(z'_t)_{t \in [0, 1]}$ satisfies the following stochastic differential equation in $\mathbb{R}^N \otimes \mathbb{R}^d$

$$dz'_t = \sum_{\ell=1}^m \nabla W_\ell(z_t) z'_t dh_t^\ell + \sum_{\ell=1}^m \nabla Z_\ell(z_t) z'_t dB_t^\ell + \nabla Z_0(z_t) z'_t dt + \sum_{\ell=1}^m Z_\ell(z_t) \otimes (v_t X_\ell(x_t)) dt. \quad (35)$$

Define processes $(m_t)_{t \in [0, 1]}$ in \mathbb{R} and $(r_t)_{t \in [0, 1]}$ in \mathbb{R}^d by

$$\begin{aligned} dm_t &= \sum_{\ell=1}^m \dot{h}_t^\ell dB_t^\ell, \quad m_0 = 0 \\ dr_t &= \sum_{\ell=1}^m v_t X_\ell(x_t) dB_t^\ell, \quad r_0 = 0. \end{aligned}$$

Write $(x'_t)_{t \in [0, 1]}$ for the derived process associated to the stochastic differential equation (31) and set

$$y_t^{(0)} = (x_{t \wedge t_1}, \dots, x_{t \wedge t_k}, x_t, v_t, m_t, r_t, x'_t).$$

Then $(y_t^{(0)})_{t \in [0, 1]}$ satisfies a stochastic differential equation of the form (33). The stopped components are obtained by multiplying the coefficients of (31) by the indicator function

$1_{[0,t_i]}(t)$. For $n \geq 0$, recursively, set $z_t^{(n)} = (y_t^{(0)}, \dots, y_t^{(n)})$, note that $(z_t^{(n)})_{t \in [0,1]}$ satisfies a stochastic differential equation of the form (33) with graded Lipschitz coefficients, define $((z^{(n)})'_t)_{t \in [0,1]}$ by solving the associated derived equation, write $(z^{(n)})'_t = ((y^{(0)})'_t, \dots, (y^{(n)})'_t)$, and set

$$y_t^{(n+1)} = (y^{(n)})'_t.$$

Where a process is presented in components, for example $y_t^{(0)}$, we will write the corresponding decomposition of its derived process in the obvious way. Thus

$$(y^{(0)})'_t = (x'_{t \wedge t_1}, \dots, x'_{t \wedge t_k}, x'_t, v'_t, m'_t, r'_t, x''_t).$$

It is straightforward to check that this notation is consistent when a given process is a component in two different autonomous processes. By a standard calculation, we have $x'_t = \sigma u_t c_t$, where c_t is the Malliavin covariance matrix

$$c_t = \sum_{\ell=1}^m \int_0^t (v_s X_\ell(x_s)) \otimes (v_s X_\ell(x_s)) ds. \quad (36)$$

It is well known that, under the bracket condition (2), the Malliavin covariance matrix c_1 is invertible and its inverse has moments of all orders. The basic form of Bismut's integration by parts formula is the identity in \mathbb{R}^d

$$\mathbb{E}(\nabla \phi(z_1) z'_1) = \mathbb{E}(\phi(z_1) r_1)$$

valid for C_b^1 functions ϕ on \mathbb{R}^N and for $(z_t)_{t \in [0,1]}$ and $(z'_t)_{t \in [0,1]}$ satisfying (33) and (35) respectively. Set

$$\tilde{x}_1 = (x_1 - \gamma_1)/\sigma, \quad \tilde{x}'_1 = x'_1/\sigma = u_1 c_1, \quad \tilde{x}''_1 = x''_1/\sigma.$$

Define random variables y in $(\mathbb{R}^d)^* \otimes (\mathbb{R}^d)^*$ and R in \mathbb{R} by

$$y = (\tilde{x}'_1)^{-1}, \quad R = (F(x_1) + \frac{1}{2}d(x, y)^2)/\sigma^2 + m_1/\sigma$$

and define y' in $(\mathbb{R}^d)^* \otimes (\mathbb{R}^d)^* \otimes \mathbb{R}^d$ and R' in \mathbb{R}^d by

$$y' = -y \tilde{x}''_1 y, \quad R' = \nabla F(x_1) x'_1 / \sigma^2 + m'_1 / \sigma.$$

Fix $n \geq 0$ and apply the integration by parts formula with $(z_t)_{t \in [0,1]}$ replaced by $(z_t^{(n)})_{t \in [0,1]}$ and with $\phi(z_1)$ replaced by $f(\tilde{x}_1) y \phi(y, z_1^{(n)}) e^{-R}$ to obtain

$$\begin{aligned} & \mathbb{E}(y \otimes (\nabla f(\tilde{x}_1) \tilde{x}'_1) \phi(y, z_1^{(n)}) e^{-R}) + \mathbb{E}(f(\tilde{x}_1) y' \phi(y, z_1^{(n)}) e^{-R}) - \mathbb{E}(f(\tilde{x}_1) \phi(y, z_1^{(n)}) (y \otimes R') e^{-R}) \\ & + \mathbb{E}(f(\tilde{x}_1) y \otimes (\nabla_y \phi(y, z_1^{(n)}) y') e^{-R}) + \mathbb{E}(f(\tilde{x}_1) y \otimes (\nabla_z \phi(y, z_1^{(n)}) (z^{(n)})'_1) e^{-R}) \\ & = \mathbb{E}(f(\tilde{x}_1) \phi(y, z_1^{(n)}) (y \otimes r_1) e^{-R}). \end{aligned}$$

We assume here that f is C_b^1 and that ϕ is polynomial in y , with coefficients C^1 in $z_1^{(n)}$ and of polynomial growth, along with their derivatives. Then all factors in the integrands of the preceding formula have moments of all orders. A straightforward approximation by C_b^1 functions then establishes the formula. Define a linear map $\tau_i : (\mathbb{R}^d)^* \otimes (\mathbb{R}^d)^* \otimes \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$\tau_i(e_j^* \otimes e_{j'}^* \otimes e_{j'}) = \delta_{ii'} \delta_{jj'}.$$

Then

$$\tau_i(y \otimes (\nabla f(\tilde{x}_1) \tilde{x}_1')) = \nabla_i f(\tilde{x}_1)$$

so

$$\mathbb{E}(\nabla_i f(\tilde{x}_1) \phi(y, z_1^{(n)}) e^{-R}) = \mathbb{E}(f(\tilde{x}_1) \nabla_i^* \phi(y, z_1^{(n+1)}) e^{-R}) \quad (37)$$

where

$$\begin{aligned} \nabla_i^* \phi(y, z_1^{(n+1)}) &= \tau_i(y \otimes r_1 + y \tilde{x}_1'' y + y \otimes R') \phi(y, z_1^{(n)}) \\ &\quad + \tau_i(y \otimes (\nabla_y \phi(y, z_1^{(n)}) y \tilde{x}_1'' y)) - \tau_i(y \otimes (\nabla_z \phi(y, z_1^{(n)})(z^{(n)})'_1)). \end{aligned}$$

Take

$$\phi_0(y, z_1^{(0)}) = G(y, z_1^{(0)}) = G(\tilde{x}) = g(\tilde{x}_{t_1}, \dots, \tilde{x}_{t_k}).$$

We see inductively that (37) is valid for $\phi_n = \nabla_{i_n}^* \dots \nabla_{i_1}^* \phi_0$ for all $n \geq 0$. So we can iterate (37) to obtain, for any multi-index $\alpha = (i_1, \dots, i_n)$

$$\mathbb{E}(\nabla^\alpha f(\tilde{x}_1) G(\tilde{x}) e^{-R}) = \mathbb{E}(f(\tilde{x}_1) (\nabla^*)^\alpha G(y, z_1^{(n)}) e^{-R}).$$

We take $f(x) = e^{i\langle \xi, x \rangle}$ to deduce that $|\xi^\alpha| |\hat{G}_\varepsilon(\xi)| \leq C_\varepsilon(\alpha, G)$ where

$$C_\varepsilon(\alpha, G) = \mathbb{E}(|(\nabla^*)^\alpha G(y, z_1^{(n)})| e^{-R}).$$

Now $y = c_1^{-1} v_1$ and $(\nabla^*)^\alpha G$ is of polynomial growth in $(y, z_1^{(n)})$. Given the estimate (26), to prove (27), it will suffice to show that, for $n = d + 1$ and for all $p < \infty$, we have

$$\sup_{\sigma \in (0,1]} \mathbb{E}(|z_1^{(n)}|^p) < \infty \quad (38)$$

and

$$\sup_{\sigma \in (0,1]} \mathbb{E}(|c_1^{-1}|^p) < \infty. \quad (39)$$

We already noted the availability of moment estimates of all orders for $z_1^{(n)}$ and that these are uniform in $\sigma \in [0, 1]$ for the components derived from $(v_t)_{t \in [0,1]}$ and $(r_t)_{t \in [0,1]}$. Recall that $\tilde{x}_t = (x_t - \gamma_t)/\sigma$ and $R = (F(x_1) + d(x, y)^2/2)/\sigma^2 + m_1/\sigma$. We will use first and second order

mean value theorems to see that there is in fact no singularity as $\sigma \rightarrow 0$ in these processes or any processes derived from them. Consider, for $\theta \in [0, 1]$, the stochastic differential equation in \mathbb{R}^d

$$dx_t(\theta) = \sum_{\ell=1}^m X_\ell(x_t(\theta))dh_t^\ell + \sum_{\ell=1}^m \theta \sigma X_\ell(x_t(\theta))dB_t^\ell + \theta^2 \sigma^2 \tilde{X}_0(x_t(\theta))dt, \quad x_0(\theta) = x.$$

There exists a unique family of strong solutions $(x_t(\theta))_{t \in [0,1]}$ which are almost surely jointly continuous in θ and t . Moreover, the following derivatives then exist for all θ and t , almost surely

$$\bar{x}_t(\theta) = \frac{1}{\sigma} \frac{\partial}{\partial \theta} x_t(\theta), \quad \hat{x}_t(\theta) = \frac{1}{\sigma^2} \left(\frac{\partial}{\partial \theta} \right)^2 x_t(\theta).$$

Moreover, the processes $(\bar{x}_t(\theta))_{t \in [0,1]}$ and $(\hat{x}_t(\theta))_{t \in [0,1]}$ satisfy

$$\begin{aligned} d\bar{x}_t(\theta) &= \sum_{\ell=1}^m \nabla X_\ell(x_t(\theta)) \bar{x}_t(\theta) dh_t^\ell + \sum_{\ell=1}^m \theta \sigma \nabla X_\ell(x_t(\theta)) \bar{x}_t(\theta) dB_t^\ell + \theta^2 \sigma^2 \nabla \tilde{X}_0(x_t(\theta)) \bar{x}_t(\theta) dt \\ &\quad + \sum_{\ell=1}^m X_\ell(x_t(\theta)) dB_t^\ell + 2\theta \sigma \tilde{X}_0(x_t(\theta)) dt, \quad \bar{x}_0(\theta) = 0 \\ d\hat{x}_t(\theta) &= \sum_{\ell=1}^m \nabla X_\ell(x_t(\theta)) \hat{x}_t(\theta) dh_t^\ell + \sum_{\ell=1}^m \theta \sigma \nabla X_\ell(x_t(\theta)) \hat{x}_t(\theta) dB_t^\ell + \theta^2 \sigma^2 \nabla \tilde{X}_0(x_t(\theta)) \hat{x}_t(\theta) dt \\ &\quad + \sum_{\ell=1}^m \nabla^2 X_\ell(x_t(\theta)) (\bar{x}_t(\theta), \bar{x}_t(\theta)) dh_t^\ell + \sum_{\ell=1}^m \theta \sigma \nabla^2 X_\ell(x_t(\theta)) (\bar{x}_t(\theta), \bar{x}_t(\theta)) dB_t^\ell \\ &\quad + \theta^2 \sigma^2 \nabla^2 \tilde{X}_0(x_t(\theta)) (\bar{x}_t(\theta), \bar{x}_t(\theta)) dt \\ &\quad + \sum_{\ell=1}^m 2 \nabla X_\ell(x_t(\theta)) \bar{x}_t(\theta) dB_t^\ell + 4\theta \sigma \nabla \tilde{X}_0(x_t(\theta)) \bar{x}_t(\theta) dt \\ &\quad + 2 \tilde{X}_0(x_t(\theta)) dt, \quad \hat{x}_0(\theta) = 0. \end{aligned}$$

Set $z_t(\theta) = (x_t(\theta), \bar{x}_t(\theta), \hat{x}_t(\theta))$. Note that the process $(z_t(\theta))_{t \in [0,1]}$ is the solution of a stochastic differential equation with graded Lipschitz coefficients, and that the coefficient bounds of the graded structure are uniform in $\theta \in [0, 1]$ and $\sigma \in [0, 1]$. Hence $(z_t(\theta))_{t \in [0,1]}$ and its derived process $(z'_t(\theta))_{t \in [0,1]}$ have moments of all orders, and these are bounded uniformly in θ and σ . Now

$$\tilde{x}_t = (x_t - \gamma_t)/\sigma = \int_0^1 \bar{x}_t(\theta) d\theta$$

and

$$R = \int_0^1 (1 - \theta) \{F'(x_1(\theta)) \hat{x}_1(\theta) + F''(x_1(\theta)) (\bar{x}_1(\theta), \bar{x}_1(\theta))\} d\theta. \quad (40)$$

We can take the derivative in (34) under the integral sign in (40) to express R' also by such an integral over θ , with integrand expressed in terms of $z_1(\theta)$ and $z'_1(\theta)$. Hence $\sup_{t \in [0,1]} |\tilde{x}_t|$, R and R' have moments of all orders which are bounded uniformly in $\sigma \in (0, 1]$. The same reasoning can be extended to conclude that all processes derived from $(\tilde{x}_t)_{t \in [0,1]}$ and R have moments of all orders bounded uniformly in $\sigma \in (0, 1]$. Alternatively, it can be observed that $(x_t, \tilde{x}'_t, v_t/\sigma, \tilde{x}''_t/\sigma, m''_t/\sigma)$ satisfies a stochastic differential equation with graded Lipschitz coefficients, with bounds uniform in $\sigma \in (0, 1]$, and from this observation we can draw the same conclusion. Hence we have shown (38).

We turn to the proof of (39). Write $(v_t(0))_{t \in [0,1]}$ for the solution to (32) when $\sigma = 0$. Recall that

$$c_1 = \sum_{\ell=1}^m \int_0^1 (v_t X_\ell(x_t)) \otimes (v_t X_\ell(x_t)) dt$$

and note that c_1 depends continuously on $(x_t, v_t)_{t \in [0,1]}$ in uniform norm. We have assumed that the deterministic Malliavin covariance matrix $c_1(0)$ is invertible. Hence there are constants $C < \infty$ and $\delta > 0$ such that $|c_1^{-1}| \leq C$ on the event

$$\Omega(\delta) = \{|x_t - \gamma_t| \leq \delta \quad \text{and} \quad |v_t - v_t(0)| \leq \delta \quad \text{for all } t \in [0, 1]\}.$$

By standard L^p estimates for stochastic differential equations, for all $p < \infty$, there is a constant $C(\delta, p) < \infty$ such that, for all $\sigma \in (0, 1]$,

$$\mathbb{P}(\Omega(\delta)^c) \leq C(\delta, p) \sigma^p.$$

On the other hand, it is shown in [3] that, for all $p < \infty$, there are constants $C(p) < \infty$ and $D \in \mathbb{N}$ such that, for all $\sigma \in (0, 1]$,

$$\|c_1^{-1}\|_p \leq C(p) \sigma^{-D}.$$

Now $|c_1^{-1}| \leq C + |c_1^{-1}| 1_{\Omega(\delta)^c}$ so, by Hölder's inequality, for all $p < \infty$ and all $\sigma \in (0, 1]$,

$$\|c_1^{-1}\|_p \leq C + C(2p) C(\delta, 2Dp)^{1/2p}.$$

Finally, consider the case where $G(\omega) = |\omega_s - \omega_t|^4$ for some $s, t \in [0, 1]$. Set $\tilde{x}_t^{(0)} = \tilde{x}_t$ and, recursively for $n \geq 0$, set $\tilde{x}_t^{(n+1)} = (\tilde{x}_t, (\tilde{x}_t^{(n)})'_t)$. Then, by standard estimates, for all $p \in [1, \infty)$, there is a constant $C < \infty$ such that, uniformly in $s, t \in [0, 1]$ and in $\sigma \in (0, 1]$, we have

$$\mathbb{E} \left(|\tilde{x}_s^{(n)} - \tilde{x}_t^{(n)}|^{4p} \right) \leq C |s - t|^{2p}.$$

The adjoint operators ∇_i^* have an explicit form written above, from which we deduce that, for all n and for $\alpha = (i_1, \dots, i_n)$, there is a random variable K_α , having moments of all orders bounded uniformly in $\sigma \in (0, 1]$, such that

$$(\nabla^*)^\alpha G(y, z_1^{(n)}) = K_\alpha |\tilde{x}_s^{(n)} - \tilde{x}_t^{(n)}|^4.$$

Hence, by Hölder's inequality, there is a constant $C_\alpha < \infty$ such that, uniformly in $s, t \in [0, 1]$ and in $\varepsilon \in (0, 1]$, we have

$$C_\varepsilon(\alpha, G) \leq C_\alpha |s - t|^2.$$

This implies (28). □

5 Localization

Having established the conclusion of Theorem 1.3 under the global condition (6), we now show that it extends to general manifolds under a condition (41) which limits the probability that the bridge strays too far from the minimal path. At the same time we obtain that this condition allows the extension of the small-time equivalent for the heat kernel. In the next section, we will find conditions on the operator which imply (41).

Theorem 5.1. *Let M be a connected C^∞ manifold. Let \mathcal{L} be a second order differential operator on M of the form (4)*

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^d a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^d b^i(x) \frac{\partial}{\partial x^i}$$

where the diffusivity a has a sub-Riemannian structure (X_1, \dots, X_m) . Suppose that $x, y \in M$ are joined by a unique minimal path γ and that (x, y) lies outside the cut locus. Suppose further that there is a chart $\phi : U_0 \rightarrow \mathbb{R}^d$ of M and a domain U in M , containing γ and compactly contained in U_0 , such that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log p(\varepsilon, x, M \setminus U, y) < -d(x, y)^2/2 \quad (41)$$

where $p(\varepsilon, x, M \setminus U, y) = p(\varepsilon, x, y) - p_U(\varepsilon, x, y)$ and $p_U(\varepsilon, x, y)$ is the Dirichlet heat kernel in U of the restriction of \mathcal{L} to U . Then, as $\varepsilon \rightarrow 0$, we have

$$p(\varepsilon, x, y) = c(x, y) \varepsilon^{-d/2} \exp\{-d(x, y)^2/(2\varepsilon)\} (1 + o(1))$$

and $\tilde{\mu}_\varepsilon^{x,y} \rightarrow \mu_\gamma$ weakly on $T_\gamma \Omega^{x,y}$.

Proof. It will suffice, and it will lighten the notation, to consider the case where U_0 is a domain both in M and in \mathbb{R}^d , where ϕ is the identity map, and where θ restricts to the identity map on U . We will show that there exist vector fields $\bar{X}_0, \bar{X}_1, \dots, \bar{X}_{m+d}$ on \mathbb{R}^d , which are bounded with bounded derivatives of all orders and such that

- (a) $(\bar{X}_1, \dots, \bar{X}_{m+d})$ is a sub-Riemannian structure on \mathbb{R}^d ,

(b) we have $a = \bar{a}$ and $\mathcal{L} = \bar{\mathcal{L}}$ on U , where

$$\bar{\mathcal{L}} = \frac{1}{2} \sum_{\ell=1}^{m+d} \bar{X}_\ell^2 + \bar{X}_0$$

and where a and \bar{a} are the diffusivities of \mathcal{L} and $\bar{\mathcal{L}}$ respectively,

(c) γ is the unique \bar{a} -minimal path in $H^{x,y}(\mathbb{R}^d)$.

From these properties, given that (x, y) lies outside the cut locus of \mathcal{L} in M , we deduce that (x, y) also lies outside the cut locus of $\bar{\mathcal{L}}$ in \mathbb{R}^d . Write \bar{p} for the heat kernel of $\bar{\mathcal{L}}$ on \mathbb{R}^d and write $\mu_\varepsilon^{x,y,\mathbb{R}^d}$ for the bridge measure on $\Omega^{x,y}(\mathbb{R}^d)$ associated to the operator $\varepsilon\bar{\mathcal{L}}$. Similarly, write $\tilde{\mu}_\varepsilon^{x,y,\mathbb{R}^d}$ for the rescaled bridge measure on $T_\gamma\Omega^{x,y}$, given by $\tilde{\mu}_\varepsilon^{x,y,\mathbb{R}^d} = \mu_\varepsilon^{x,y,\mathbb{R}^d} \circ \bar{\sigma}_\varepsilon^{-1}$, where $\bar{\sigma}_\varepsilon(\omega)_t = (\omega_t - \gamma_t)/\sqrt{\varepsilon}$. Then, as we showed in Section 4, as $\varepsilon \rightarrow 0$, we have

$$\bar{p}(\varepsilon, x, y) = c(x, y)\varepsilon^{-d/2} \exp\{-d(x, y)^2/(2\varepsilon)\}(1 + o(1))$$

and $\mu_\varepsilon^{x,y,\mathbb{R}^d} \rightarrow \delta_\gamma$ weakly on $\Omega^{x,y}(\mathbb{R}^d)$ and $\tilde{\mu}_\varepsilon^{x,y,\mathbb{R}^d} \rightarrow \mu_\gamma$ weakly on $T_\gamma\Omega^{x,y}$.

Write $\mu_\varepsilon^{x,y,U}$ for the diffusion bridge measure on $\Omega^{x,y}(U)$ associated to the restriction of the operator $\varepsilon\mathcal{L}$ to U . Then, for all measurable sets A in $\Omega^{x,y}(M)$, we have

$$p(\varepsilon, x, y)\mu_\varepsilon^{x,y}(A) = p_U(\varepsilon, x, y)\mu_\varepsilon^{x,y,U}(A \cap \Omega^{x,y}(U)) + p(\varepsilon, x, y)\mu_\varepsilon^{x,y}(A \setminus \Omega^{x,y}(U))$$

and

$$\bar{p}(\varepsilon, x, y)\mu_\varepsilon^{x,y,\mathbb{R}^d}(A \cap \Omega^{x,y}(U)) = p_U(\varepsilon, x, y)\mu_\varepsilon^{x,y,U}(A \cap \Omega^{x,y}(U)).$$

For any bounded measurable set B in $T_\gamma\Omega^{x,y}$ and for $\varepsilon > 0$ sufficiently small, we have $\sigma_\varepsilon^{-1}(B) = \bar{\sigma}_\varepsilon^{-1}(B) \subseteq \Omega^{x,y}(U)$, so

$$p(\varepsilon, x, y)\tilde{\mu}_\varepsilon^{x,y}(B) = \bar{p}(\varepsilon, x, y)\tilde{\mu}_\varepsilon^{x,y,\mathbb{R}^d}(B).$$

Now $\mu_\varepsilon^{x,y,\mathbb{R}^d}(\Omega^{x,y}(U)) \rightarrow 1$ so, taking $A = \Omega^{x,y}(M)$, we find

$$p_U(\varepsilon, x, y) = c(x, y)\varepsilon^{-d/2} \exp\{-d(x, y)^2/(2\varepsilon)\}(1 + o(1))$$

and, using (41),

$$p(\varepsilon, x, y) = p_U(\varepsilon, x, y) + p(\varepsilon, x, M \setminus U, y) = c(x, y)\varepsilon^{-d/2} \exp\{-d(x, y)^2/(2\varepsilon)\}(1 + o(1)).$$

Moreover, for all $B \in T_\gamma\Omega^{x,y}$, we then have $\tilde{\mu}_\varepsilon^{x,y}(B) = \tilde{\mu}_\varepsilon^{x,y,\mathbb{R}^d}(B)(1 + o(1))$, so the claimed weak limit for $\tilde{\mu}_\varepsilon^{x,y}$ follows from that for $\tilde{\mu}_\varepsilon^{x,y,\mathbb{R}^d}$.

It remains to show the existence of vector fields $\bar{X}_0, \bar{X}_1, \dots, \bar{X}_{m+d}$ with the claimed properties. Fix an open set U_1 such that U is compactly contained in U_1 and U_1 is compactly

contained in U_0 . There exists a C^∞ function χ on \mathbb{R}^d such that $1_U \leq \chi \leq 1$ and which is uniformly positive on U_1 and such that $\{\chi > 0\}$ is compactly contained in U_0 . Write X_0 for the vector field on M given by (5). For $\ell = 0, 1, \dots, m$, set $\bar{X}_\ell(z) = \chi(z)X_\ell(z)$ for $z \in U_0$ and set $\bar{X}_\ell(z) = 0$ for $z \in \mathbb{R}^d \setminus U_0$. Then \bar{X}_ℓ is a bounded vector field on \mathbb{R}^d with bounded derivatives of all orders. Define $\tilde{a} = \sum_{\ell=1}^m \bar{X}_\ell \otimes \bar{X}_\ell$ and write \tilde{I} for the associated energy function. Then $\tilde{I}(\omega) \geq I(\omega)$ for all $\omega \in H^{x,y}(U_0)$ with equality if ω is contained in U . Moreover, by choosing χ sufficiently small near ∂U_1 , we can and do ensure that, for all $z \in \partial U_1$ and all $\omega \in H^{x,z}(U_0)$, we have $\tilde{I}(\omega) > I(\gamma)$. Choose now another C^∞ function $\bar{\chi}$ on \mathbb{R}^d such that $1_{U_1} \leq 1 - \bar{\chi} \leq 1_{U_0}$ and such that $\chi + \bar{\chi}$ is everywhere positive. For $i = 1, \dots, d$, set $\bar{X}_{m+i}(z) = \bar{\chi}(z)e_i$, where e_1, \dots, e_d is the standard basis in \mathbb{R}^d . Then the brackets of $\bar{X}_1, \dots, \bar{X}_m$ span \mathbb{R}^d on $\{\chi > 0\}$, while $\bar{X}_{m+1}, \dots, \bar{X}_{m+d}$ themselves span \mathbb{R}^d on $\{\bar{\chi} > 0\}$. Hence (a) holds. Also, (b) holds because $\bar{X}_\ell = X_\ell$ on U for $\ell = 0, 1, \dots, m$ and $\bar{X}_\ell = 0$ on U for $\ell = m+1, \dots, m+d$. Now, if $\omega \in H^{x,y}(\mathbb{R}^d)$ is contained in U_1 , then $\bar{I}(\omega) = \tilde{I}(\omega) \geq I(\omega)$, so $\bar{I}(\omega) \geq I(\gamma)$ with equality only if $\omega = \gamma$. On the other hand, if ω is not contained in U_1 , set $\tau = \inf\{t \in [0, 1] : \omega_t \in \partial U_1\}$ and set $\tilde{\omega}_t = \omega_{\tau t}$. Then $\tilde{\omega} \in H^{x,z}(U_0)$, where $z = \omega_\tau \in \partial U_1$. So $\bar{I}(\omega) \geq \tilde{I}(\tilde{\omega}) > I(\gamma)$. Hence (c) holds. \square

6 Heat kernel upper bounds

Let M be a connected C^∞ manifold. Consider a second order differential operator \mathcal{L} on M of the form (8)

$$\mathcal{L}f = \frac{1}{2} \operatorname{div}(a \nabla f) + a(\beta, \nabla f)$$

where the diffusivity a has a sub-Riemannian structure (X_1, \dots, X_m) , where the divergence is understood with respect to a positive C^∞ locally invariant measure ν and where β is a C^∞ 1-form satisfying the sector condition (7). Recall the definition (3) of the sub-Riemannian distance. For a closed subset K of M and $x, y \in M \setminus K$, consider the distance from x to y through K , given by

$$d(x, K, y) = \inf\{\sqrt{I(\omega)} : \omega \in H^{x,y}, \omega_t \in K \text{ for some } t \in [0, 1]\}.$$

The heat kernel for diffusion from x to y through K is defined by

$$p(t, x, K, y) = p(t, x, y) - p_{M \setminus K}(t, x, y)$$

where $p_{M \setminus K}$ is the Dirichlet heat kernel of \mathcal{L} in $M \setminus K$. We will prove in this section the following upper bound.⁹

⁹In the case where a is everywhere positive-definite, there is an alternative argument of Hsu [12] for heat kernel upper bounds in incomplete manifolds, relying on estimates of Azencott [4], which does not require the condition (7). Hsu instead requires that $d(x, y) \leq d(x, \infty) + d(\infty, y)$. The same paper [12] also contains an example where heat flow does not localize near the shortest path.

Theorem 6.1. *For all $x, y \in M$ and all closed subsets K of M , we have*

$$\limsup_{t \rightarrow 0} t \log p(t, x, K, y) \leq -d(x, K, y)^2/2.$$

Before giving the proof, we show how this estimate allows us to prove Theorems 1.1, 1.2 and 1.3.

Proof of Theorem 1.1. The upper bound in (10) is the simple case $K = M$ of Theorem 6.1, while the lower bound follows by a standard argument from Léandre's lower bound in \mathbb{R}^d .

We turn to the proof of (11). Consider the operator

$$\tilde{\mathcal{L}} = \mathcal{L} + \frac{1}{2}(\partial/\partial\tau)^2$$

on $\tilde{M} = M \times \mathbb{R}$, where τ denotes the coordinate in \mathbb{R} . Then the diffusivity of $\tilde{\mathcal{L}}$ also has a sub-Riemannian structure. Set $\tilde{x} = (x, 0)$ and $\tilde{y} = (y, 1)$, and define

$$\tilde{K} = \tilde{M} \setminus \tilde{U}, \quad \tilde{U} = \{(\gamma_t, \sigma_t) : (\gamma, \sigma) \in \tilde{\Gamma}(\delta), t \in [0, 1]\}$$

where

$$\tilde{\Gamma}(\delta) = \left\{ (\gamma, \sigma) \in H^{x,y} \times H^{0,1}(\mathbb{R}) : I(\gamma) + \int_0^1 \dot{\sigma}_t^2 dt < d(x, y)^2 + 1 + \delta \right\}.$$

Then \tilde{K} is closed in \tilde{M} . Write $\beta_\varepsilon^{0,1}$ for the law on $\Omega^{0,1}(\mathbb{R})$ of a Brownian bridge from 0 to 1 of speed ε . Then, with obvious notation,

$$\tilde{p}(t, \tilde{x}, \tilde{y}) = p(t, x, y) \frac{1}{\sqrt{2\pi}} e^{-1/(2t)}, \quad \tilde{\mu}_\varepsilon^{\tilde{x}, \tilde{y}}(d\omega, d\tau) = \mu_\varepsilon^{x,y}(d\omega) \beta_\varepsilon^{0,1}(d\tau).$$

By Theorem 6.1, we have

$$\limsup_{t \rightarrow 0} t \log \tilde{p}(t, \tilde{x}, \tilde{K}, \tilde{y}) \leq -\tilde{d}(\tilde{x}, \tilde{K}, \tilde{y})^2/2 = -(d(x, y)^2 + 1 + \delta)/2$$

so

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \tilde{\mu}_\varepsilon^{\tilde{x}, \tilde{y}}(\{(\omega, \tau) : (\omega_t, \tau_t) \in \tilde{K} \text{ for some } t \in [0, 1]\}) \\ \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \tilde{p}(\varepsilon, \tilde{x}, \tilde{K}, \tilde{y}) - \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \tilde{p}(\varepsilon, \tilde{x}, \tilde{y}) \leq -\delta/2 \end{aligned} \quad (42)$$

where we have used the lower bound from (10). By standard estimates, we also have

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log \beta_\varepsilon^{0,1}(\{\tau : |\tau_t - t| \geq \sqrt{\delta}/2 \text{ for some } t \in [0, 1]\}) = -\delta/2. \quad (43)$$

Suppose then that $\omega \in \Omega^{x,y}$ and $\tau \in \Omega^{0,1}(\mathbb{R})$ satisfy $(\omega_t, \tau_t) \in \tilde{U}$ and $|\tau_t - t| < \sqrt{\delta}/2$ for all $t \in [0, 1]$. Then, for each $t \in [0, 1]$, there exist $s \in [0, 1]$ and $\gamma \in H^{x,y}$ and $\sigma \in H^{0,1}(\mathbb{R})$ such that

$$\omega_t = \gamma_s, \quad \tau_t = \sigma_s, \quad I(\gamma) < d(x, y)^2 + \delta, \quad \int_0^1 \dot{\sigma}_r^2 dr < 1 + \delta.$$

Then $|\sigma_s - s| \leq \sqrt{\delta}/2$ so $|t - s| \leq \sqrt{\delta}$ and so

$$d(\omega_t, \Gamma_t(\delta))^2 \leq d(\omega_t, \gamma_t)^2 = d(\gamma_s, \gamma_t)^2 \leq |t - s|I(\gamma) \leq \delta^{1/2}(d(x, y)^2 + \delta).$$

The estimates (42) and (43) thus imply (11). \square

Proof of Theorems 1.2 and 1.3. The hypothesis that $\gamma \in H^{x,y}$ is strongly minimal implies that, for any domain U containing γ , we have $d(x, M \setminus U, y) > d(x, y)$. We can choose a chart U_0 in M and a domain U containing γ such that U is compactly contained in U_0 . We suppose that \mathcal{L} has the form (8) and satisfies the sector condition (7). Then Theorem 6.1 applies to show that

$$\limsup_{t \rightarrow 0} t \log p(t, x, M \setminus U, y) \leq -d(x, M \setminus U, y)^2/2 < -d(x, y)^2/2$$

and Theorem 5.1 then gives the desired conclusions. \square

In proving Theorem 6.1, we will use the following estimate of Nagel, Stein & Wainger [23] for the sub-Riemannian distance. There is a covering of M by relatively compact charts $\phi : U \rightarrow \mathbb{R}^d$ such that, for some constants $\alpha(U) > 0$ and $C(U) < \infty$, for all $x, y \in U$, we have

$$C^{-1}|\phi(x) - \phi(y)| \leq d(x, y) \leq C|\phi(x) - \phi(y)|^\alpha.$$

It follows that, for any positive C^∞ measure ν on M , we can choose $C(U)$ so that moreover, for all $x \in U$ and all $r \in (0, \infty)$ such that $B(x, r) = \{y \in M : d(x, y) < r\} \subseteq U$, we have

$$C^{-1}r^{d/\alpha} \leq \nu(B(x, r)) \leq Cr^d \quad (44)$$

In particular, $C(U)$ may be chosen so that the following local volume-doubling inequality holds: for all $x \in U$ and all $r \in (0, \infty)$ such that $B(x, 2r) \subseteq U$, we have

$$\nu(B(x, 2r)) \leq C\nu(B(x, r)). \quad (45)$$

We will also need the local Poincaré inequality proved by Jerison [14], which also relies on the fact that a has a sub-Riemannian structure. There is a covering of M by open sets U such that, for some constant $C(U) < \infty$, for all $x \in U$ and all $r \in (0, \infty)$ such that $B(x, 2r) \subseteq U$, for all $f \in C_c^\infty(M)$, we have

$$\int_{B(x, r)} |f - \langle f \rangle_{B(x, r)}|^2 d\nu \leq Cr^2 \int_{B(x, 2r)} a(\nabla f, \nabla f) d\nu \quad (46)$$

where $\langle f \rangle_B = \int_B f d\nu / \nu(B)$ is the average value of f on B . The sector condition (7) gives us uniform parabolicity in the sense of Sturm, so (45) and (46) allow us to apply [28, Theorem 2.1] to obtain the following parabolic mean-value estimate. For simplicity, we will state a simple version adequate for our needs. Let $x \in M$ and let $r_0 \in (0, \infty)$. Let D be a domain in M which compactly contains the ball $B(x, r_0)$. Suppose that u is a non-negative weak solution of the heat equation $(\partial/\partial t)u_t = \mathcal{L}u_t$ on $(0, \infty) \times D$. Then there is constant $C(D, r_0) < \infty$ such that, for all $t \in (0, \infty)$ and all $r \in (0, r_0]$ with $r^2 \leq t/2$, we have

$$u_t(x)^2 \leq C \int_{t-r^2}^t \int_{B(x,r)} u_s^2 d\nu ds. \quad (47)$$

See Lierl and Saloff-Coste [19, Theorem 2.6] for an alternative proof, which clarifies some points in the non-symmetric case.

The sub-Riemannian distance has a dual formulation, proved in Jerison & Sanchez-Calle [15], which we adapt to the case of a general sub-Riemannian manifold, without completeness, and to the distance through K .

Proposition 6.2. *For all $x, y \in M$ and any closed subset K of M , we have*

$$d(x, K, y) = \sup\{w^+(y) - w^-(x) : w^-, w^+ \in \mathcal{F} \text{ with } w^+ = w^- \text{ on } K\}$$

where \mathcal{F} is the set of all locally Lipschitz functions w on M such that $a(\nabla w, \nabla w) \leq 1$ almost everywhere.

Proof. Denote the right hand side by $\delta(x, K, y)$ for now. Suppose that $\omega \in H^{x,y}$ has driving path ξ and that $\omega_t \in K$. Let $w^-, w^+ \in \mathcal{F}$, with $w^+ = w^-$ on K . It will suffice to consider the case where ω is simple, when there exists a relatively compact chart U for M containing ω . Then, given $\varepsilon > 0$, since a is continuous, we can find C^∞ functions f^-, f^+ on U such that $|f^\pm(z) - w^\pm(z)| \leq \varepsilon$ and $a(\nabla f^\pm, \nabla f^\pm)(z) \leq 1 + \varepsilon$ for all $z \in U$. Then

$$w^+(y) - w^-(x) = w^+(y) - w^+(\omega_t) + w^-(\omega_t) - w^-(x) \leq f^+(y) - f^+(\omega_t) + f^-(\omega_t) - f^-(x) + 4\varepsilon$$

and

$$\begin{aligned} & f^+(y) - f^+(\omega_t) + f^-(\omega_t) - f^-(x) \\ &= \int_0^t \langle \nabla f^-(\omega_s), \dot{\omega}_s \rangle ds + \int_t^1 \langle \nabla f^+(\omega_s), \dot{\omega}_s \rangle ds \\ &= \int_0^t \langle \nabla f^-(\omega_s), a(\omega_s) \xi_s \rangle ds + \int_t^1 \langle \nabla f^+(\omega_s), a(\omega_s) \xi_s \rangle ds \\ &\leq \left(\int_0^t a(\nabla f^-, \nabla f^-)(\omega_s) ds + \int_t^1 a(\nabla f^+, \nabla f^+)(\omega_s) ds \right)^{1/2} \left(\int_0^1 a(\xi_s, \xi_s) ds \right)^{1/2} \\ &\leq (1 + \varepsilon) \sqrt{I(\omega)}. \end{aligned}$$

Hence $w^+(y) - w^-(x) \leq \sqrt{I(\omega)}$. On taking the supremum over w^\pm and the infimum over ω , we deduce that $\delta(x, K, y) \leq d(x, K, y)$.

Now we prove the reverse inequality. Choose a C^∞ quadratic form α on T^*M which is everywhere positive-definite. Consider the energy function I_α and the distance functions d_α and δ_α obtained by replacing a by $a + \alpha$ in the definitions of d and δ . Set $w^+(z) = d_\alpha(x, K, z)$ and $w^-(z) = d_\alpha(x, z)$. Note that $w^+ = w^-$ on K . Since $a + \alpha$ is positive-definite, the functions w^- and w^+ are locally Lipschitz, and their weak gradients ∇w^\pm satisfy $(a + \alpha)(\nabla w^\pm, \nabla w^\pm) \leq 1$ almost everywhere. Hence

$$d_\alpha(x, K, y) = w^+(y) - w^-(x) \leq \delta_\alpha(x, K, y) \leq \delta(x, K, y)$$

and we can complete the proof by finding, for each $\varepsilon \in (0, 1]$, a choice of α so that $d(x, K, y) \leq (1 + \varepsilon)(d_\alpha(x, K, y) + \varepsilon) + \varepsilon$.

Choose $\phi^\alpha \in \Omega^{x,y}$, passing through K and such that $\sqrt{I_\alpha(\phi^\alpha)} \leq d_\alpha(x, K, y) + \varepsilon$. There exist $p \in \mathbb{N}$ and C^∞ vector fields Y_1, \dots, Y_p on M such that $\alpha = \sum_{i=1}^p Y_i \otimes Y_i$. Then there exist $h \in H^0(\mathbb{R}^m)$ and $k \in H^0(\mathbb{R}^p)$ such that ϕ^α has weak derivative¹⁰

$$\dot{\phi}_t^\alpha = \sum_{\ell=1}^m X_\ell(\phi_t^\alpha) \dot{h}_t^\ell + \sum_{i=1}^p Y_i(\phi_t^\alpha) \dot{k}_t^i.$$

By reparametrizing ϕ^α if necessary, we may assume that $|\dot{h}_t|^2 + |\dot{k}_t|^2 = I_\alpha(\phi^\alpha)$ for almost all t . It will be convenient to assume that $d(x, \infty) \geq 1/2$, which we may do without loss of generality by a scaling argument. Define $S_{-1} = S_0 = \emptyset$ and consider for $n \geq 1$ the compact set

$$S_n = \{z \in M : d(z, x) \leq 2^n \text{ and } d(z, \infty) \geq 2^{-n}\}.$$

Define, recursively, sequences of times t_0, t_1, \dots, t_J , points x_1, \dots, x_J and y_1, \dots, y_J , and positive integers n_1, \dots, n_J as follows. Set $t_0 = 0$ and $n_1 = 1$. For $j \geq 1$, set $x_j = \phi_{t_{j-1}}^\alpha$, define $(\phi_t^j)_{t \geq t_{j-1}}$ by the differential equation

$$\dot{\phi}_t^j = \sum_{\ell=1}^m X_\ell(\phi_t^j) \dot{h}_t^\ell, \quad \phi_{t_{j-1}}^j = x_j$$

and set $y_j = \phi_{t_j}^j$, where

$$t_j = \inf\{t \geq t_{j-1} : \phi_t^j \in \partial S_{n_j-1} \cup \partial S_{n_j+1}\} \wedge 1.$$

If $t_j < 1$, then set

$$n_{j+1} = \begin{cases} n_j - 1, & \text{if } y_j \in \partial S_{n_j-1}, \\ n_j + 1, & \text{if } y_j \in \partial S_{n_j+1}. \end{cases}$$

¹⁰See the discussion preceding Proposition 7.1.

Thus $y_j \in \partial S_{n_{j+1}}$. If $t_j = 1$, then set $J = j$ and stop. We will see below that the recursion does stop. Note that, for all j and all $s, t \in [t_{j-1}, t_j]$ with $s \leq t$, we have

$$d(\phi_s^j, \phi_t^j) \leq (t - s) \|\dot{h}\|_\infty.$$

Note also that $x_1 = \phi_0^\alpha = x \in S_{n_1} \subseteq S_{n_1+1} \setminus S_{n_1-1}$ and $\phi_{t_{j-1}}^\alpha = \phi_{t_{j-1}}^j = x_j$ for all j . Suppose, inductively for $j \geq 1$, that $x_j \in S_{n_j+1} \setminus S_{n_j-1}$. Then ϕ_t^j remains in this set for $t \in [t_{j-1}, t_j]$ and, by choosing α sufficiently small on $S_{n_j+2} \setminus S_{n_j-2}$, we can ensure that

$$d(\phi_t^j, \phi_t^\alpha) \leq 2^{-2n_j-2}\varepsilon \leq 2^{-n_{j+1}-2}, \quad t \in [t_{j-1}, t_j].$$

Then, in particular, we have $d(y_j, x_{j+1}) \leq 2^{-n_{j+1}-2}$. Now $y_j \in \partial S_{n_{j+1}}$ so this implies that $x_{j+1} \in S_{n_{j+1}+1} \setminus S_{n_{j+1}-1}$ and the induction proceeds. Also, since $y_{j+1} \in \partial S_{n_{j+2}}$, we have $d(y_j, y_{j+1}) \geq 2^{-n_{j+1}-1}$ and so $d(x_{j+1}, y_{j+1}) \geq 2^{-n_{j+1}-2}$. Certainly $d(x_1, y_1) \geq 2^{-n_1-2}$ so, for all $j \geq 1$, we have $d(x_j, y_j) \geq 2^{-n_j-2}$. Hence

$$(t_j - t_{j-1}) \|\dot{h}\|_\infty \geq 2^{-n_j-2}$$

and so, for all $n \geq 1$, we have $n_j = n$ at most $2^{n+2} \|\dot{h}\|_\infty$ times. In particular, the recursion must stop, or $n_j \rightarrow \infty$ as $j \rightarrow \infty$ forcing $\phi_{t_j}^\alpha$ to leave all compact sets, which is impossible. Hence

$$\sum_{j=1}^{J-1} d(y_j, x_{j+1}) + d(y_J, y) \leq \sum_{n=1}^{\infty} 2^{n+2} \|\dot{h}\|_\infty \cdot 2^{-2n-2}\varepsilon = \varepsilon \|\dot{h}\|_\infty.$$

Also, there exist j and t so that $t \in [t_{j-1}, t_j]$ and $\phi_t^\alpha \in K$, and then

$$d(x_j, K, y_j) \leq d(x_j, \phi_t^j) + 2d(\phi_t^j, \phi_t^\alpha) + d(\phi_t^j, y_j) \leq (t_j - t_{j-1}) \|\dot{h}\|_\infty + \varepsilon$$

while, for $k \neq j$, we have the estimate

$$d(x_k, y_k) \leq (t_k - t_{k-1}) \|\dot{h}\|_\infty.$$

We finally combine these estimates, using the triangle inequality, to obtain

$$d(x, K, y) \leq (1 + \varepsilon) \|\dot{h}\|_\infty + \varepsilon \leq (1 + \varepsilon)(d_\alpha(x, K, y) + \varepsilon) + \varepsilon.$$

□

Proof of Theorem 6.1. We will show that the argument used in [26, Theorem 1.2], for the case where a is positive-definite and $\beta = 0$, generalizes to the present context¹¹. Consider the set

¹¹The idea is to combine a standard argument for heat kernel upper bounds with a reflection trick. In terms of Markov processes, we give a random sign to each excursion of the diffusion process into D , viewing it as taking values in D^- or D^+ . Then a generalization of the classical reflection principle for Brownian motion allows to express the density for paths from x to y via K in terms of this enhanced process. In fact the heat kernel \tilde{p} for this process may be written in terms of p and p_D , and we find it technically simpler to define \tilde{p} in those terms, rather than set up the enhanced process.

$\tilde{M} = M^- \cup M^+$, where $M^\pm = K \cup D^\pm$ and D^-, D^+ are disjoint copies of $D = M \setminus K$. Write π for the obvious projection $\tilde{M} \rightarrow M$. For functions f defined on M , we will write f also for the function $f \circ \pi$ on \tilde{M} . Thus we will sometimes consider a as a quadratic form on T^*D^\pm and β as a 1-form on D^\pm . Define a measure $\tilde{\nu}$ on \tilde{M} by

$$\tilde{\nu}(A) = \nu(A \cap K) + \frac{1}{2}\nu(\pi(A \cap D^-)) + \frac{1}{2}\nu(\pi(A \cap D^+)).$$

Note that $\nu = \tilde{\nu} \circ \pi^{-1}$. Now define

$$\tilde{p}(t, x, y) = \begin{cases} p(t, x, y) + p_D(t, x, y), & \text{if } x, y \in D^\pm, \\ p(t, x, y) - p_D(t, x, y), & \text{if } x \in D^\pm \text{ and } y \in D^\mp, \\ p(t, x, y), & \text{if } x \in K \text{ or } y \in K. \end{cases}$$

Given functions $f^-, f^+ \in C_c^\infty(M)$ with $f^- = f^+$ on K , write f for the function on \tilde{M} such that $f = f^\pm \circ \pi$ on M^\pm , and set $\bar{f} = (f^- + f^+)/2$ and $f^D = (f^+ - f^-)/2$. Let $\phi^\pm \in C_c^\infty(M)$ with $\phi^- = \phi^+$ on K and define ϕ on \tilde{M} and $\bar{\phi}$ and ϕ^D on M similarly. Define for $t \in (0, \infty)$ functions u_t on \tilde{M} , \bar{u}_t on M and u_t^D on D by

$$u_t(x) = \int_{\tilde{M}} \tilde{p}(t, x, y) f(y) \tilde{\nu}(dy)$$

and

$$\bar{u}_t(x) = \int_M p(t, x, y) \bar{f}(y) \nu(dy), \quad u_t^D(x) = \int_M p_D(t, x, y) f^D(y) \nu(dy).$$

Then \bar{u}_t and u_t^D solve the heat equation with Dirichlet boundary conditions in M and D respectively. It is straightforward to check that $u_t = u_t^\pm \circ \pi$ on M^\pm , where $u_t^\pm = \bar{u}_t \pm u_t^D$ and we extend u_t^D by 0 on K . Hence

$$\int_{\tilde{M}} \phi u_t d\tilde{\nu} = \int_M \bar{\phi} \bar{u}_t d\nu + \int_D \phi^D u_t^D d\nu$$

and so

$$\begin{aligned} \frac{d}{dt} \int_{\tilde{M}} \phi u_t d\tilde{\nu} &= \frac{d}{dt} \int_M \bar{\phi} \bar{u}_t d\nu + \frac{d}{dt} \int_D \phi^D u_t^D d\nu \\ &= -\frac{1}{2} \int_M a(\nabla \bar{\phi}, \nabla \bar{u}_t) d\nu + \int_M a(\bar{\phi} \beta, \nabla \bar{u}_t) d\nu - \frac{1}{2} \int_D a(\nabla \phi^D, \nabla u_t^D) d\nu + \int_D a(\phi^D \beta, \nabla u_t^D) d\nu \\ &= -\frac{1}{2} \int_{\tilde{M}} a(\nabla \phi, \nabla u_t) d\tilde{\nu} + \int_{\tilde{M}} a(\phi \beta, \nabla u_t) d\tilde{\nu}. \end{aligned} \tag{48}$$

Fix $x, y \in D$ and $r_0 \in (0, \infty)$ so that the balls $B(x, r_0)$ and $B(y, r_0)$ are compactly contained in D . Write x^- and y^+ for the unique points in D^- and D^+ respectively such that $\pi(x^-) = x$ and $\pi(y^+) = y$. Fix $r \in (0, r_0]$ and set

$$B^- = \{z \in D^- : \pi(z) \in B(x, r)\}, \quad B^+ = \{z \in D^+ : \pi(z) \in B(y, r)\}.$$

Consider the set \mathcal{F}_r of pairs of bounded locally Lipschitz functions (w^-, w^+) on M such that w^- is constant on $B(x, r)$, w^+ is constant on $B(y, r)$, $w^- = w^+$ on K and $a(\nabla w^\pm, \nabla w^\pm) \leq 1$ almost everywhere. Set

$$d_r(x, K, y) = \sup\{w^+(y) - w^-(x) : (w^-, w^+) \in \mathcal{F}_r\}.$$

Fix $(w^-, w^+) \in \mathcal{F}_r$ and define a function w on \tilde{M} by setting $w = w^\pm \circ \pi$ on M^\pm . Take $f^- = 0$ and choose $f^+ \geq 0$ supported on $B(y, r)$ and such that $\int_M f^2 d\nu = 2$. Then $\int_{\tilde{M}} \tilde{f}^2 d\tilde{\nu} = 1$. Fix $\delta \in [0, \infty)$ and $\theta \in \mathbb{R}$ and set $\psi = \theta w$. We deduce from (48) by a standard argument that

$$\begin{aligned} \frac{d}{dt} \int_{\tilde{M}} (e^\psi u_t)^2 d\tilde{\nu} &= - \int_{\tilde{M}} a(\nabla(e^{2\psi} u_t), \nabla u_t) d\tilde{\nu} + 2 \int_{\tilde{M}} a(\beta e^{2\psi} u_t, \nabla u_t) d\tilde{\nu} \\ &= - \int_{\tilde{M}} a(\nabla u_t, \nabla u_t) e^{2\psi} d\tilde{\nu} + 2 \int_{\tilde{M}} a((\beta - \nabla \psi) u_t, \nabla u_t) e^{2\psi} d\tilde{\nu} \\ &\leq \int_{\tilde{M}} a(\beta - \nabla \psi, \beta - \nabla \psi) (e^\psi u_t)^2 d\tilde{\nu} \leq \lambda \int_{\tilde{M}} (e^\psi u_t)^2 d\tilde{\nu} \end{aligned}$$

where $\lambda = \|a(\beta - \nabla \psi, \beta - \nabla \psi)\|_\infty \leq (1 + \delta)\theta^2 + (1 + 1/\delta)\|a(\beta, \beta)\|_\infty$. So, by Gronwall, we have

$$\int_{\tilde{M}} (e^\psi u_t)^2 d\tilde{\nu} \leq e^{\lambda t} \int_{\tilde{M}} (e^\psi \tilde{f})^2 d\tilde{\nu} = e^{\lambda t + 2\psi(y^+)}$$

and so

$$\int_{B^-} u_t^2 d\tilde{\nu} \leq e^{2\theta(w^+(y) - w^-(x)) + \lambda t}.$$

Moreover, since $u_t^- \geq 0$ and $(\partial/\partial t)u_t^- = \mathcal{L}u_t^-$ on $(0, \infty) \times D$, by the parabolic mean-value estimate, there is a constant $C(D, r_0) < \infty$ such that, for all $t \in (0, \infty)$ and all $r \in (0, r_0]$ with $r^2 \leq t/2$, we have

$$u_t(x^-)^2 \leq C \int_{t-r^2}^t \int_{B^-} u_s^2 d\tilde{\nu} ds \leq C \nu(B(x, r))^{-1} e^{2\theta(w^+(y) - w^-(x)) + \lambda t}.$$

Set $v_t(z) = p(t, x, K, z)$, then $(\partial/\partial t)v_t = \hat{\mathcal{L}}v_t$ on $(0, \infty) \times D$, where $\hat{\mathcal{L}}f = \frac{1}{2} \operatorname{div}(a \nabla f) - a(\beta, \nabla f)$. So, by the parabolic mean-value estimate again, we can choose $C(D, r_0)$ so that

$$p(t, x, K, y)^2 \leq C \int_{t-r^2}^t \int_{B(y, r)} p(s, x, K, z)^2 \nu(dz) ds = C \int_{t-r^2}^t \int_{B^+} \tilde{p}(s, x^-, z)^2 \tilde{\nu}(dz) ds.$$

Assume now that $r^2 \leq t/4$. For each $s \in [t - r^2, t]$, we may take $f^+ = cp(s, x, K, \cdot)1_{B(y, r)}$, where c is chosen so that $\int_{\tilde{M}} f^2 d\tilde{\nu} = 1$. For this choice of f^+ , we have

$$u_s(x^-)^2 = \int_{B^+} \tilde{p}(s, x^-, z)^2 \tilde{\nu}(dz).$$

Hence

$$\begin{aligned} p(t, x, K, y)^2 &\leq C\nu(B(y, r))^{-1} \int_{t-r^2}^t u_s(x^-)^2 ds \\ &\leq C^2\nu(B(x, r))^{-1}\nu(B(y, r))^{-1} e^{2\theta(w^+(y) - w^-(x)) + \lambda t} \end{aligned}$$

from which we obtain, on optimizing over θ and w ,

$$p(t, x, K, y) \leq C\nu(B(x, r))^{-1/2}\nu(B(y, r))^{-1/2} \exp \left\{ -\frac{d_r(x, K, y)^2}{2(1+\delta)t} + \left(1 + \frac{1}{\delta}\right) \frac{\|a(\beta, \beta)\|_\infty t}{2} \right\}.$$

For $(w^-, w^+) \in \mathcal{F}_0$ and $r \in (0, r_0]$, set $\bar{w}_x = \sup_{z \in B(x, r)} w^-(z)$ and $\bar{w}_y = \inf_{z \in B(x, r)} w^+(z)$ and suppose that $\bar{w}_x \leq \bar{w}_y$. We can define $(w_r^-, w_r^+) \in \mathcal{F}_r$ by setting $w_r^-(z) = \bar{w}_x \vee w(z) \wedge \bar{w}_y$ and $w_r^+(z) = \bar{w}_x \vee w(z) \wedge \bar{w}_y$. Note that $|w^-(z) - w^-(x)| \leq r$ for all $z \in B(x, r)$ and $|w^+(z) - w^+(y)| \leq r$ for all $z \in B(y, r)$, so $w^+(y) - w^-(x) \leq w_r^+(y) - w_r^-(x) + 2r$. On optimizing over w , we see that

$$d(x, K, y) \leq d_r(x, K, y) + 2r.$$

Hence, using the volume estimate (44), we finally obtain, for all $t \in (0, 4r_0^2 \vee 1]$

$$p(t, x, K, y) \leq Ct^{-d/(2\alpha)}(1 + d(x, K, y)^2/t)^{d/(2\alpha)} \exp \left\{ -\frac{d(x, K, y)^2}{2(1+\delta)t} + \left(1 + \frac{1}{\delta}\right) \frac{\|a(\beta, \beta)\|_\infty t}{2} \right\}$$

which is certainly sufficient to imply the claimed asymptotics. \square

7 Second variation of the energy in a sub-Riemannian manifold of non-constant rank

Let M be a connected C^∞ manifold of dimension d and let a be a C^∞ non-negative quadratic form on T^*M having a sub-Riemannian structure, as in Section 1. In the section and the next, we assume that M is complete for the sub-Riemannian distance. Since all questions which we address concern properties determined for a finite energy path by any neighbourhood of that path, this assumption of completeness results in no essential loss of generality. Recall,

for a continuous path ω in M , the notions of energy and driving path defined in Section 1. In Section 2, we reviewed the notion of cut locus and defined the Gaussian measure μ_γ in terms of the bicharacteristic flow. Extrapolating from the Riemannian case, we might hope to characterize these objects instead in terms of the energy function I . This will be done in Section 8. In the present section, in preparation, we show that the set $H^{x,y}$ of finite-energy paths has, at suitably selected paths ω , a well-defined set of tangent directions $T_\omega H^{x,y}$, and that $T_\omega H^{x,y}$ has the structure of a Hilbert space. Then we show that, when ω is minimal and ξ is a driving path for ω , the energy has a well-defined second variation Q_ξ in a dense set of tangent directions, which allows us to define a continuous non-negative quadratic form Q_ξ on $T_\omega H^{x,y}$. Finally, we show that Q_ξ is minimized over ξ by a unique driving path λ , which is in fact a bicharacteristic. Given a sub-Riemannian structure X for a , there are related constructions on the set of control paths, which are well known, evidently depending on the choice of X . We emphasise that $T_\omega H^{x,y}$ and $Q = Q_\lambda$ depend on a alone.

Two sub-Riemannian structures (X_1, \dots, X_m) and (Y_1, \dots, Y_n) are equivalent (see [1]) if $X_\ell = \sum_{k=1}^n f_{\ell k} Y_k$ and $Y_k = \sum_{\ell=1}^m g_{k\ell} X_\ell$ for all ℓ and k , for some C^∞ functions $f_{\ell k}$ and $g_{k\ell}$ on M . If a has constant rank, then all sub-Riemannian structures for a are equivalent. However, this is not true in general, as the following example shows. In \mathbb{R}^2 , take

$$\begin{aligned} X_1(x, y) = Y_1(x, y) &= y \frac{\partial}{\partial x}, & X_2(x, y) = Y_2(x, y) &= \frac{\partial}{\partial y} \\ X_3(x, y) = \operatorname{sgn}(x) Y_3(x, y) &= e^{-1/|x|} \frac{\partial}{\partial x}. \end{aligned} \tag{49}$$

Then

$$\sum_{\ell=1}^3 X_\ell \otimes X_\ell = \sum_{\ell=1}^3 Y_\ell \otimes Y_\ell$$

but X and Y define inequivalent sub-Riemannian structures on \mathbb{R}^2 . Thus, in the non-constant rank case, we cannot establish that an object is intrinsic to a by showing it is intrinsic to an equivalence class of sub-Riemannian structures. Instead, our approach will be to work directly from a , using a sub-Riemannian structure for existence but not uniqueness.

Recall that H^x denotes the set of finite-energy paths starting at x and $H^{x,y}$ denotes the set of such paths terminating at y . Also, $H^0(\mathbb{R}^m)$ denotes the space of absolutely continuous paths $h : [0, 1] \rightarrow \mathbb{R}^m$ starting from 0 such that

$$\|h\|^2 = \int_0^1 |\dot{h}_t|^2 dt < \infty.$$

We fix a sub-Riemannian structure $X = (X_1, \dots, X_m)$ for a and use this to construct some associated objects. We will make clear which objects depend on the choice of X and which

do not. Given $\omega \in H^x$, we define $h(\omega) \in H^0(\mathbb{R}^m)$ by

$$\dot{h}_t(\omega) = X(\omega_t)^* \xi_t$$

where ξ is any driving path for ω . Then $h(\omega)$ does not depend on the choice of ξ . For $x \in M$ and $h \in H^0(\mathbb{R}^m)$, write $\phi(x, h)$ for the solution $(\phi_t)_{t \in [0,1]}$ of the differential equation

$$\dot{\phi}_t = \sum_{\ell=1}^m X_\ell(\phi_t) \dot{h}_t^\ell, \quad \phi_0 = x.$$

Denote by $p_X(x)$ the orthogonal projection $\mathbb{R}^m \rightarrow (\ker X(x))^\perp$. For $x \in M$ and $h, k \in H^0(\mathbb{R}^m)$, define $\pi = \pi(x, h)k \in H^0(\mathbb{R}^m)$ by

$$\dot{\pi}_t = p_X(\phi_t(x, h)) \dot{k}_t.$$

The following parametrization of H^x by $H^0(\mathbb{R}^m)$ using X is well known.

Proposition 7.1. *Let $x \in M$ and let $\gamma \in H^x$. Then $\gamma = \phi(x, h(\gamma))$. Moreover, for all $h \in H^0(\mathbb{R}^m)$, if we set $\omega = \phi(x, h)$ and $\pi = \pi(x, h)h$, then $\omega \in H^x$ and $\pi = h(\omega)$ and $I(\omega) = \|\pi\|^2 \leq \|h\|^2$.*

Proof. Since $(\ker X(x))^\perp = \text{im } X(x)^*$, there is a measurable map $\xi : [0, 1] \rightarrow T^*M$ over ω such that $\dot{\pi}_t = X(\omega_t)^* \xi_t$. Then $\dot{\omega}_t = X(\omega_t)X(\omega_t)^* \xi_t = a(\omega_t) \xi_t$ and

$$I(\omega) = \int_0^1 \langle \xi_t, a(\omega_t) \xi_t \rangle dt = \int_0^1 |X(\omega_t)^* \xi_t|^2 dt = \|\pi\|^2.$$

We leave the remaining details to the reader. □

Let ω be a finite-energy path and fix a chart along ω . We say that a driving path ξ for ω is *tame* if

$$\int_0^1 |\xi_t|^2 dt < \infty$$

where we have used the Euclidean norm in the chart. This condition does not depend on the choice of chart. We say that ω is *tame*¹² if it has a tame driving path.

Let ω be a tame path in H^x . Fix a chart along ω and a tame driving path ξ for ω . We will define a space $T_\omega H^x$ of *finite-energy variations* of ω which, for now, may appear to

¹²It is straightforward to see that, if the diffusivity a has constant rank, then every finite-energy path is tame. On the other hand, consider on \mathbb{R}^2 the vector fields $X_1 = \partial/\partial x_1, X_2 = x_1 \partial/\partial x_2$ and the path $\omega_t = (t, t^2/2)$ for $t \in [0, 1]$. Since X_1, X_2 span \mathbb{R}^2 except on $\{x_1 = 0\}$, the only driving path for ω is $\xi_t = dx_1 + (1/t)dx_2$. Then $\langle \xi_t, a(\omega_t) \xi_t \rangle = 2$ for all t so ω has finite energy, but ξ is not tame.

depend on the choice of chart and of ξ . Denote by $T_\omega H^x$ the set of absolutely continuous maps $v : [0, 1] \rightarrow TM$ over ω such that

$$\dot{v}_t = (\nabla_{v_t} a)(\omega_t) \xi_t + a(\omega_t) \eta_t, \quad v_0 = 0 \quad (50)$$

for some measurable path $\eta : [0, 1] \rightarrow T^*M$ over ω , with

$$\|v\|_\xi^2 := \int_0^1 \langle \eta_t, a(\omega_t) \eta_t \rangle dt < \infty.$$

Note that, in (50), the meaning of the derivatives \dot{v}_t and $\nabla_{v_t} a$ depends on the choice of chart. Note also that, since ξ is tame, by Gronwall's lemma, there is a constant $C < \infty$ such that $\|v\|_\infty \leq C \|v\|_\xi$ for all $v \in T_\omega H^x$. If η can be chosen in (50) so that, in addition,

$$\int_0^1 |\eta_t|^2 dt < \infty$$

then we will call v a *tame finite-energy variation*. Write $\tilde{T}_\omega H^x$ for the set of tame finite-energy variations of ω . Set $T_\omega H^{x,y} = \{v \in T_\omega H^x : v_1 = 0\}$ and $\tilde{T}_\omega H^{x,y} = \{v \in \tilde{T}_\omega H^x : v_1 = 0\}$.

Proposition 7.2. *Let $\omega \in H^x$ be tame, with tame driving path ξ . Then $\tilde{T}_\omega H^x$ is dense in $T_\omega H^x$. Moreover $\tilde{T}_\omega H^{x,y}$ is dense in $T_\omega H^{x,y}$.*

Proof. For $\varepsilon > 0$, set $\eta_t^\varepsilon = (\sqrt{a(\omega_t)} + \varepsilon I)^{-1} \sqrt{a(\omega_t)} \eta_t$ and define v^ε by

$$\dot{v}_t^\varepsilon = (\nabla_{v_t^\varepsilon} a)(\omega_t) \xi_t + a(\omega_t) \eta_t^\varepsilon, \quad v_0^\varepsilon = 0.$$

Note that $\sqrt{a(\omega_t)}(\eta_t^\varepsilon - \eta_t) \rightarrow 0$ for all $t \in [0, 1]$ and

$$\langle \eta_t^\varepsilon - \eta_t, a(\omega_t)(\eta_t^\varepsilon - \eta_t) \rangle = \varepsilon^2 |\eta_t^\varepsilon|^2 \leq \langle \eta_t^\varepsilon, a(\omega_t) \eta_t^\varepsilon \rangle + \varepsilon^2 |\eta_t^\varepsilon|^2 \leq \langle \eta_t, a(\omega_t) \eta_t \rangle.$$

Hence

$$\int_0^1 \langle \eta_t^\varepsilon, a(\omega_t) \eta_t^\varepsilon \rangle dt + \varepsilon^2 \int_0^1 |\eta_t^\varepsilon|^2 dt \leq \int_0^1 \langle \eta_t, a(\omega_t) \eta_t \rangle dt < \infty$$

so $v^\varepsilon \in \tilde{T}_\omega H^x$ for all ε , and by dominated convergence, as $\varepsilon \rightarrow 0$,

$$\|v^\varepsilon - v\|_\xi^2 = \int_0^1 \langle \eta_t^\varepsilon - \eta_t, a(\omega_t)(\eta_t^\varepsilon - \eta_t) \rangle dt \rightarrow 0.$$

We have proved the first assertion.

Set $V = \{v_1 : v \in T_\omega H^x\}$. There is a finite set $E \subseteq T_\omega H^x$ such that $\{e_1 : e \in E\}$ is a basis for V . Given $\delta > 0$, we can find for each $e \in E$ a tame $\tilde{e} \in \tilde{T}_\omega H^x$ such that $\|e - \tilde{e}\|_\xi \leq \delta$.

Then $|e_1 - \tilde{e}_1| \leq C\delta$. We can therefore choose δ sufficiently small that $\{\tilde{e}_1 : e \in E\}$ remains a basis for V . Set

$$A := \sup \left\{ \left\| \sum_{e \in E} \alpha_e \tilde{e} \right\|_{\xi} : \alpha_e \in \mathbb{R}, \left| \sum_{e \in E} \alpha_e \tilde{e}_1 \right| = 1 \right\}.$$

Then $A < \infty$. Given $v \in T_{\omega}H^{x,y}$ and $\varepsilon > 0$, there exists $v' \in \tilde{T}_{\omega}H^x$ such that $\|v - v'\|_{\xi} < \varepsilon/(1 + CA)$. Then $v'_1 \in V$ and $|v'_1| \leq C\varepsilon/(1 + CA)$. We can write $v'_1 = \sum_{e \in E} \alpha_e \tilde{e}_1$ for some $\alpha_e \in \mathbb{R}$. Set

$$\tilde{v} = v' - \sum_{e \in E} \alpha_e \tilde{e}$$

then $\tilde{v} \in \tilde{T}_{\omega}H^{x,y}$ and

$$\|v - \tilde{v}\|_{\xi} \leq \|v - v'\|_{\xi} + \left\| \sum_{e \in E} \alpha_e \tilde{e} \right\|_{\xi} \leq \varepsilon$$

which proves the second assertion. \square

By standard arguments, the map $\phi : M \times H^0(\mathbb{R}^m) \rightarrow H$ is differentiable, in both arguments. Fix $x \in M$ and $h \in H^0(\mathbb{R}^m)$. Set $\omega_t = \phi_t(x, h)$ and define

$$u_t = \phi_t^*(x, h) = \frac{\partial}{\partial x} \phi_t(x, h) \in T_{\omega_t}M \otimes T_x^*M.$$

For $k \in H^0(\mathbb{R}^m)$, define $v = v(k)$ by

$$v_t = \frac{\partial}{\partial h} \phi_t(x, h)k \in T_{\omega_t}M.$$

Then u and v satisfy the differential equations

$$\begin{aligned} \dot{u}_t &= \sum_{\ell=1}^m \nabla X_{\ell}(\omega_t) u_t \dot{h}_t^{\ell}, \quad u_0 = I, \\ \dot{v}_t &= \sum_{\ell=1}^m \nabla X_{\ell}(\omega_t) v_t \dot{h}_t^{\ell} + X_{\ell}(\omega_t) \dot{k}_t^{\ell}, \quad v_0 = 0 \end{aligned} \tag{51}$$

where the derivatives \dot{u}_t , \dot{v}_t and ∇X_{ℓ} are understood in the chosen chart. Then, by the variation of constants formula,

$$v_t = \sum_{\ell=1}^m u_t \int_0^t u_s^{-1} X_{\ell}(\omega_s) \dot{k}_s^{\ell} ds. \tag{52}$$

Proposition 7.3. *Let $\omega \in H^x$ be tame and let $k \in H^0(\mathbb{R}^m)$. Fix a chart along ω and a tame driving path ξ for ω , and write $T_\omega H^x$ for the associated space of finite-energy variations. Set*

$$v(k) = \frac{\partial}{\partial h} \phi(x, h(\omega)) k.$$

Then $v(k) \in T_\omega H^x$. On the other hand, for any $v \in T_\omega H^x$, we can define $k(\xi, v) \in H^0(\mathbb{R}^m)$ by

$$\dot{k}_t^\ell(\xi, v) = \langle \xi_t, \nabla X_\ell(\omega_t) v_t \rangle + \langle \eta_t, X_\ell(\omega_t) \rangle$$

*where $\eta : [0, 1] \rightarrow T^*M$ is a measurable path over ω satisfying (50). Then $k(\xi, v)$ does not depend on the choice of η , and*

$$v = \frac{\partial}{\partial h} \phi(x, h(\omega)) k(\xi, v).$$

Moreover there is a constant $C < \infty$, depending only on $I(\omega)$, $\int_0^1 |\xi_t|^2 dt$ and a uniform bound for X and ∇X along ω , such that

$$\|v(k)\|_\xi \leq C \|k\|, \quad \|k(\xi, v)\| \leq C \|v\|_\xi.$$

Moreover, if $v = v(k)$, then $\pi(x, h(\omega))(k(\xi, v) - k) = 0$.

Proof. Write $h = h(\omega)$. Then $v(k)$ satisfies (51) so, by Gronwall's lemma, there is a constant $C < \infty$ such that $\|v(k)\|_\infty \leq C \|k\|$. Here and below, C is understood to have the dependence claimed in the statement. Since ξ is tame, we can define $g(k) \in H^0(\mathbb{R}^m)$ by

$$\dot{g}_t^\ell(k) = \langle \xi_t, \nabla X_\ell(\omega_t) v_t(k) \rangle.$$

There is a constant $C < \infty$ such that $\|g(k)\| \leq C \|v(k)\|_\infty$. We can find a measurable map $\eta : [0, 1] \rightarrow T^*M$ over ω , such that

$$a(\omega_t) \eta_t = \sum_{\ell=1}^m X_\ell(\omega_t) (\dot{k}_t^\ell - \dot{g}_t^\ell(k)). \quad (53)$$

Note that

$$\nabla a(x) = \sum_{\ell=1}^m \nabla X_\ell(x) X_\ell(x)^* + X_\ell(x) \nabla X_\ell(x)^*$$

so $v(k)$ satisfies

$$\dot{v}_t(k) = (\nabla_{v_t(k)} a)(\omega_t) \xi_t + a(\omega_t) \eta_t, \quad v_0(k) = 0$$

and, moreover

$$\|v(k)\|_\xi^2 = \int_0^1 \langle \eta_t, a(\omega_t) \eta_t \rangle dt = \|\pi(x, h)(k - g)\|^2 \leq C\|k\|^2.$$

Hence $v(k) \in T_\omega H^x$ and $\|v(k)\|_\xi \leq C\|k\|$.

On the other hand, for $v \in T_\omega H^x$, we see from (50) that there is a constant $C < \infty$ such that $\|v\|_\infty \leq C\|v\|_\xi$, so $\|k(\xi, v)\| \leq C\|v\|_\xi$. Moreover, we can write (50) in the form

$$\dot{v}_t = \sum_{\ell=1}^m \nabla X_\ell(\omega_t) v_t \dot{h}_t^\ell + X_\ell(\omega_t) \dot{k}_t^\ell(\xi, v)$$

so $v = v(k(\xi, v))$, as claimed. Finally, if $v = v(k)$, then $v(k(\xi, v) - k) = 0$, so, from (52),

$$\sum_{\ell=1}^m X_\ell(\omega_t) (\dot{k}_t^\ell(\xi, v) - \dot{k}_t^\ell) = 0$$

for almost all t , and so $\pi(x, h)(k(\xi, v) - k) = 0$. \square

Fix ω and set $\tilde{H} = \{k \in H^0(\mathbb{R}^m) : \pi(x, h(\omega))k = k\}$. Then \tilde{H} is a closed subspace of the Hilbert space $H^0(\mathbb{R}^m)$. Proposition 7.3 shows that, for any choice of chart along ω and any choice of tame driving path ξ for ω , the map $k \mapsto v(k)$ is a linear isomorphism $\tilde{H} \rightarrow T_\omega H^x$, which is bounded with bounded inverse, when $T_\omega H^x$ is given the norm $\|\cdot\|_\xi$. Hence, the space $T_\omega H^x$ does not depend on the choice of chart and driving path. It clearly does not depend either on any choice of sub-Riemannian structure X . Moreover, the norms $\|\cdot\|_\xi$ are all equivalent, and all make $T_\omega H^x$ into a Hilbert space.

A (minimizing) finite-energy path $\omega \in H^{x,y}$ is said to be *regular* if the linear map

$$\frac{\partial}{\partial h} \phi_1(x, h(\omega)) : H^0(\mathbb{R}^m) \rightarrow T_y M$$

is onto. By (52), this is equivalent to Bismut's condition that the *deterministic Malliavin covariance matrix*

$$C_1(\omega) = \sum_{\ell=1}^m \int_0^1 (u_t^{-1} X_\ell(\omega_t)) \otimes (u_t^{-1} X_\ell(\omega_t)) dt = \int_0^1 u_t^{-1} a(\omega_t) (u_t^{-1})^* dt \quad (54)$$

is invertible. In particular, when a is positive-definite, every $\omega \in H^{x,y}$ is regular. In general, these conditions may depend on the choice of sub-Riemannian structure. By Proposition 7.3, a tame path $\omega \in H^{x,y}$ is regular if and only

$$\{v_1 : v \in T_\omega H^x\} = T_y M$$

so for tame paths the notion of regularity depends only on a and not on the choice of sub-Riemannian structure. It is straightforward to see that ω is regular if and only if its time-reversal is regular. In [7, pp. 22–24], Bismut gives an argument which shows that invertibility of $C_1(\omega)$ is intrinsic to the equivalence class of the sub-Riemannian structure X . As the example (49) shows, this is not the same as being intrinsic to a .

Consider a regular tame finite-energy path $\omega \in H^{x,y}$. We use the sub-Riemannian structure X to define $h(\omega) \in H^0(\mathbb{R}^m)$ as above and we write K for the kernel of the linear map $(\partial/\partial h)\phi_1(x, h(\omega)) : H^0(\mathbb{R}^m) \rightarrow T_y M$. Then

$$K^\perp = \{k \in H^0(\mathbb{R}^m) : \dot{k}_t^\ell = \langle \eta_0, u_t^{-1} X_\ell(\omega_t) \rangle, \eta_0 \in T_x^* M\}$$

and $(\partial/\partial h)\phi_1(x, h(\omega))|_{K^\perp}$ is invertible. By the implicit function theorem in Hilbert space, there exist $\delta > 0$ and a C^∞ map $\theta : K \rightarrow H$ such that, for all $k \in K$, we have

$$\phi_1(x, h(\omega) + k + \theta(k)) = y$$

and such that, for all $k \in K$ and all $k' \in K^\perp$ with $\|k + k'\| < \delta$, we have $\theta(k) \in K^\perp$ and

$$\phi_1(x, h + k + k') = y \quad \text{only if} \quad k' = \theta(k).$$

Note that $\theta(0) = 0$. For $k \in K$ and ε sufficiently small, we have

$$\phi_1(x, h + \varepsilon k + \theta(\varepsilon k)) = y.$$

On differentiating in ε at 0, we obtain $(\partial/\partial h)\phi_1(x, h)(k + \theta'(0)k) = 0$, so $\theta'(0)k \in K$. Since θ takes values in K^\perp , we deduce that $\theta'(0) = 0$. On taking the second derivative, we obtain

$$\frac{\partial^2}{\partial h^2}\phi_1(x, h)(k, k) + \frac{\partial}{\partial h}\phi_1(x, h)\theta''(0)(k, k) = 0. \quad (55)$$

Since $\theta''(0)(k, k) \in K^\perp$, this equation determines $\theta''(0)$.

We note the following useful identity. Let $k \in K^\perp$ and let $k' \in H^0(\mathbb{R}^m)$. Then $\dot{k}_t^\ell = \langle \eta_0, u_t^{-1} X_\ell(\omega_t) \rangle$ for some $\eta_0 \in T_x^* M$. So

$$\langle k, k' \rangle = \int_0^1 \langle \eta_0, u_t^{-1} X_\ell(\omega_t) \rangle (\dot{k}'_t)^\ell dt = \langle \eta_0, u_1^{-1} v_1(k') \rangle. \quad (56)$$

The key arguments and computations, for a sub-Riemannian structure X , in the next two results are due to Bismut [7, Theorem 1.17 and Theorem 1.24]. Our new contribution is to construct objects and show results which do not depend on the choice of X .

Proposition 7.4. *Let $\omega \in H^{x,y}$ be tame and regular and let ξ be a tame driving path for ω . Let $v \in \tilde{T}_\omega H^{x,y}$ be a tame finite-energy variation. There exists a measurable map*

$$(\varepsilon, t) \mapsto \xi_t^\varepsilon : (-1, 1) \times [0, 1] \rightarrow T^*M$$

such that $\xi^0 = \xi$ and

- (i) $\omega^\varepsilon = \pi \xi^\varepsilon \in H^{x,y}$ for all ε , with $\dot{\omega}_t^\varepsilon = a(\omega_t^\varepsilon) \xi_t^\varepsilon$,
- (ii) in any chart along ω , there is a constant $C < \infty$ such that, for all $\varepsilon \in (-1, 1)$,

$$\sup_{t \in [0,1]} |\omega_t^\varepsilon - \omega_t - \varepsilon v_t| \leq C\varepsilon^2$$

and, writing $\xi^\varepsilon = (\omega^\varepsilon, p^\varepsilon)$ and $\eta = (\omega, q)$,

$$\int_0^1 \langle p_t^\varepsilon - p_t - \varepsilon q_t, a(\omega_t^\varepsilon)(p_t^\varepsilon - p_t - \varepsilon q_t) \rangle dt \leq C\varepsilon^4.$$

Moreover, for any such map $\varepsilon \mapsto \xi^\varepsilon$, the map $\varepsilon \mapsto I(\omega^\varepsilon)$ is differentiable at $\varepsilon = 0$, and is twice differentiable at $\varepsilon = 0$ if ω is minimal. Define

$$L(v) = \left. \frac{\partial}{\partial \varepsilon} \right|_{\varepsilon=0} I(\omega^\varepsilon)$$

then L extends uniquely to a continuous linear form on $T_\omega H^{x,y}$. In the case where ω is minimal, define

$$Q_\xi(v) = \left. \frac{1}{2} \frac{\partial^2}{\partial \varepsilon^2} \right|_{\varepsilon=0} I(\omega^\varepsilon) \tag{57}$$

then Q_ξ extends uniquely to a continuous quadratic form on $T_\omega H^{x,y}$. Finally, given a sub-Riemannian structure X for a , for $h = h(\omega)$ and $k = k(\xi, v)$, we have

$$L(v) = 2\langle h, k \rangle$$

and, when ω is minimal, then $h \in K^\perp$ and

$$Q_\xi(v) = \|k\|^2 + \langle h, \theta''(0)(k, k) \rangle.$$

Proof. Fix a sub-Riemannian structure X for a . We will use X to construct a map ξ^ε having the claimed properties. Set $h = h(\omega)$ and $k = k(\xi, v)$. Since $v_1 = 0$, we have $k \in K$. For $\varepsilon \in (-1, 1)$, set $h^\varepsilon = h + \varepsilon k + \theta(\varepsilon k)$ and set $\omega^\varepsilon = \phi(x, h^\varepsilon)$. Then $\omega^\varepsilon \in H^{x,y}$ for all ε . Moreover, the map $\varepsilon \mapsto \omega^\varepsilon : (-1, 1) \rightarrow \Omega$ is C^1 and there is a constant $C < \infty$ such that

$$\sup_{t \in [0,1]} |\omega_t^\varepsilon - \omega_t - \varepsilon v_t| \leq C\varepsilon^2.$$

Moreover, fixing a chart along ω and writing $\xi = (\omega, p)$ and $\eta = (\omega, q)$, we have

$$\dot{\omega}_t^\varepsilon - a(\omega_t^\varepsilon)(p_t + \varepsilon q_t) = X_\ell(\omega_t^\varepsilon) \dot{g}_t^{\varepsilon, \ell}$$

where

$$\begin{aligned} \dot{g}_t^{\varepsilon, \ell} &= \dot{h}_t^\ell + \varepsilon \dot{k}_t^\ell + \dot{\theta}_t^\ell(\varepsilon k) - \langle p_t + \varepsilon q_t, X_\ell(\omega_t^\varepsilon) \rangle \\ &= -\langle p_t, X_\ell(\omega_t^\varepsilon) - X_\ell(\omega_t) - \varepsilon \nabla X_\ell(\omega_t) v_t \rangle - \varepsilon \langle q_t, X_\ell(\omega_t^\varepsilon) - X_\ell(\omega_t) \rangle + \dot{\theta}_t^\ell(\varepsilon k). \end{aligned}$$

Since ξ and η are tame, there is a constant $C < \infty$ such that, for all ε , we have

$$\int_0^1 |\dot{g}_t^\varepsilon|^2 dt \leq C\varepsilon^4$$

so we can find a measurable map r_t^ε such that

$$X_\ell(\omega_t^\varepsilon) \dot{g}_t^{\varepsilon, \ell} = a(\omega_t^\varepsilon) r_t^\varepsilon$$

and

$$\int_0^1 \langle r_t^\varepsilon, a(\omega_t^\varepsilon) r_t^\varepsilon \rangle dt \leq C\varepsilon^4.$$

If we now set $p_t^\varepsilon = p_t + \varepsilon q_t + r_t^\varepsilon$, then $\xi_t^\varepsilon = (\omega_t^\varepsilon, p_t^\varepsilon)$ has the required properties.

Suppose now, more generally, that $\xi^\varepsilon = (\omega^\varepsilon, p^\varepsilon)$ and $\eta = (\omega, q)$ are maps having the properties described in the statement. Define $h^\varepsilon \in H^0(\mathbb{R}^m)$ by

$$\dot{h}_t^{\varepsilon, \ell} = \langle p_t^\varepsilon, X_\ell(\omega_t^\varepsilon) \rangle.$$

Then

$$\begin{aligned} \dot{h}_t^{\varepsilon, \ell} - \dot{h}_t^\ell - \varepsilon \dot{k}_t^\ell \\ = \langle p_t^\varepsilon - p_t - \varepsilon q_t, X_\ell(\omega_t^\varepsilon) \rangle + \varepsilon \langle q_t, X_\ell(\omega_t^\varepsilon) - X_\ell(\omega_t) \rangle + \langle p_t, X_\ell(\omega_t^\varepsilon) - X_\ell(\omega_t) - \varepsilon \nabla X_\ell(\omega_t) v_t \rangle \end{aligned}$$

so, since ξ and η are tame, there is a constant $C < \infty$ such that, for all ε , we have

$$\|h^\varepsilon - h - \varepsilon k\| \leq C\varepsilon^2.$$

Write k^ε for the orthogonal projection of $\varepsilon^{-1}(h^\varepsilon - h)$ onto K . Since $k \in K$, we have $\|k^\varepsilon - k\| \leq \|\varepsilon^{-1}(h^\varepsilon - h) - k\| \leq C\varepsilon$. Since $\phi_1(x, h^\varepsilon) = y$, we have, for ε sufficiently small,

$$h^\varepsilon = h + \varepsilon k^\varepsilon + \theta(\varepsilon k^\varepsilon).$$

Hence, as $\varepsilon \rightarrow 0$,

$$I(\omega^\varepsilon) - I(\omega) = \|h + \varepsilon k^\varepsilon + \theta(\varepsilon k^\varepsilon)\|^2 - \|h\|^2 = 2\varepsilon \langle h, k \rangle + O(\varepsilon^2).$$

Hence $\varepsilon \mapsto I(\omega^\varepsilon)$ is differentiable at $\varepsilon = 0$ with derivative $L(v) = 2\langle h, k \rangle$. By Proposition 7.3, we have $\|k(\xi, v)\| \leq C\|v\|_\xi$, so L is continuous on $\tilde{T}_\omega H^{x,y}$, which is dense in $T_\omega H^{x,y}$, so L extends uniquely to $T_\omega H^{x,y}$.

Now, if ω is minimal, we must have $L(v) = 0$ for all $v \in \tilde{T}_\omega H^{x,y}$ and hence for all $v \in T_\omega H^{x,y}$. So, for any $k' \in K$, we have

$$\langle h, k' \rangle = \langle h, k(\xi, v(k')) \rangle = L(v(k')) = 0$$

and so $h \in K^\perp$. Then, for ω^ε as above, as $\varepsilon \rightarrow 0$, we have

$$I(\omega^\varepsilon) - I(\omega) = \varepsilon^2 \|k^\varepsilon\|^2 + 2\langle h, \theta(\varepsilon k^\varepsilon) \rangle + \|\theta(\varepsilon k^\varepsilon)\|^2 = \varepsilon^2 \{\|k\|^2 + \langle h, \theta''(0)(k, k) \rangle\} + O(\varepsilon^3).$$

This shows that $\varepsilon \mapsto I(\omega^\varepsilon)$ is twice differentiable at $\varepsilon = 0$ with the claimed second derivative, which is then continuous on $\tilde{T}_\omega H^{x,y}$ and extends uniquely to $T_\omega H^{x,y}$. \square

Note that, since $L(v)$ and $Q_\xi(v)$ can be computed either by choosing a suitable family of paths $\varepsilon \mapsto \xi^\varepsilon$ or by choosing a sub-Riemannian structure X , they depend on neither choice. For $\gamma \in H^{x,y}$ tame, regular and minimal, we define a continuous quadratic form Q on $T_\gamma H^{x,y}$ by

$$Q(v) = \inf_\xi Q_\xi(v)$$

where the infimum is taken over all tame driving paths ξ .

Proposition 7.5. *Let $\gamma \in H^{x,y}$ be tame, regular and minimal. Then there exists a unique tame driving path λ such that $Q = Q_\lambda$. The path λ is a bicharacteristic and is the only bicharacteristic which is a driving path for γ . Moreover, given a sub-Riemannian structure X for a , we have*

$$Q_\xi(v) = Q(v) + \|k(\xi, v) - k(\lambda, v)\|^2, \quad Q(v) = q(k(\lambda, v))$$

where q is the quadratic form on $K = \ker(\partial/\partial h)\phi_1(x, h(\gamma))$ given by

$$q(k) = \|k\|^2 - \left\langle \lambda_1, \frac{\partial^2}{\partial h^2} \phi_1(x, h(\gamma))(k, k) \right\rangle.$$

Proof. Choose a sub-Riemannian structure X for a and define as above

$$h = h(\gamma), \quad u_t = \frac{\partial}{\partial x} \phi_t(x, h), \quad K = \ker \frac{\partial}{\partial h} \phi_1(x, h).$$

Then, by Proposition 7.4, we have $h \in K^\perp$, so there exists a unique $\lambda_0 \in T_x^* M$ such that

$$\dot{h}_t^\ell = \langle \lambda_t, X_\ell(\gamma_t) \rangle$$

for almost all t , where $\lambda_t = (u_t^{-1})^* \lambda_0$. Fix a chart along γ and write $\lambda_t = (\gamma_t, p_t)$. Then

$$\dot{\gamma}_t = \sum_{\ell=1}^m X_\ell(\gamma_t) \dot{h}_t^\ell, \quad \dot{p}_t = - \sum_{\ell=1}^m \langle p_t, \nabla X_\ell(\gamma_t) \rangle \dot{h}_t^\ell$$

so λ is a bicharacteristic. Suppose, on the other hand, that $\xi = (\gamma, q)$ is a bicharacteristic over γ and define $k \in H^0(\mathbb{R}^m)$ by $\dot{k}_t^\ell = \langle \xi_t, X_\ell(\gamma_t) \rangle$. Then $\gamma = \phi(x, k)$ and $I(\gamma) = \|k\|^2$ and

$$\dot{q}_t = - \sum_{\ell=1}^m \langle q_t, \nabla X_\ell(\gamma_t) \rangle \dot{k}_t^\ell$$

so $k = h(\gamma)$ and $q = p$, proving uniqueness.

Write h for $h(\gamma)$ and recall that $\dot{h}_t^\ell = \langle \lambda_0, u_t^{-1} X_\ell(\gamma_t) \rangle$ and $\lambda_1 = (u_1^{-1})^* \lambda_0$. Then, by (55) and (56), we have

$$q(k) = \|k\|^2 + \langle h, \theta''(0)(k, k) \rangle. \quad (58)$$

So $q(k(\xi, v)) = Q_\xi(v)$ by Proposition 7.4. Take now $v = v(k)$ and set $k' = k - k(\lambda, v)$. Then $(\partial/\partial h)\phi(x, h)k' = 0$, so $\sum_{\ell=1}^m X_\ell(\gamma_t)(\dot{k}'_t)^\ell = 0$ for almost all t , and so $\phi(x, h + \varepsilon k') = \gamma$ for all $\varepsilon \in \mathbb{R}$. Hence $q(k') = \|k'\|^2$ and it will suffice to show that $q(k(\lambda, v), k') = 0$. Recall that $\dot{k}_t^\ell(\lambda, v) = \langle \lambda_t, \nabla X_\ell(\gamma_t) v_t \rangle + \langle \eta_t, X_\ell(\gamma_t) \rangle$. We differentiate the identity

$$\left(\frac{\partial}{\partial x} \phi_1(x, h) \right)^{-1} \frac{\partial}{\partial h} \phi_1(x, h) k' = \sum_{\ell=1}^m \int_0^1 \left(\frac{\partial}{\partial x} \phi_t(x, h) \right)^{-1} X_\ell(\phi_t(x, h)) (\dot{k}'_t)^\ell dt \quad (59)$$

in h , in the direction $k(\lambda, v)$, to obtain

$$u_1^{-1} \frac{\partial^2}{\partial h^2} \phi_1(x, h)(k(\lambda, v), k') = \sum_{\ell=1}^m \int_0^1 u_t^{-1} \nabla X_\ell(\gamma_t) v_t (\dot{k}'_t)^\ell dt.$$

Hence

$$\left\langle \lambda_1, \frac{\partial^2}{\partial h^2} \phi_1(x, h)(k(\lambda, v), k') \right\rangle = \sum_{\ell=1}^m \int_0^1 \langle \lambda_t, \nabla X_\ell(\gamma_t) v_t \rangle (\dot{k}'_t)^\ell dt = \langle k(\lambda, v), k' \rangle.$$

This shows that $q(k(\lambda, v), k') = 0$ as required. \square

Note that our extra condition of tameness, under which regularity of finite-energy paths is intrinsic to a , does not exclude any normal minimal paths as these all have smooth driving paths.

8 Characterization of the cut locus and fluctuation measure by the energy function

Throughout this section, γ is a regular tame minimal path in $H^{x,y}$, λ is the unique bicharacteristic projecting to γ and Q is the quadratic form on $T_\gamma H^{x,y}$ defined in the preceding section. Fix $s \in [0, 1]$ and $\eta_s \in T_{\gamma_s}^* M$. For $t \in [s, 1]$, define $J_{ts} : T_{\gamma_s}^* M \rightarrow T_{\gamma_t} M$ by

$$J_{ts}\eta_s = \left. \frac{\partial}{\partial \varepsilon} \right|_{\varepsilon=0} \pi \psi_{t-s}(\lambda_s + \varepsilon \eta_s)$$

where $(\psi_t)_{t \in \mathbb{R}}$ is the bicharacteristic flow. We already defined $J_t = J_{t0}$ by equation (15) above. Set $v_t = J_{ts}\eta_s$. Given a sub-Riemannian structure X for a , the following equations determine $\phi_{ts}(x, h)$ and \dot{h}_t , for $t \in [s, 1]$ and x near γ_t , as functions of $\lambda_s \in T_{\gamma_s}^* M$

$$\phi_{ss}(x, h) = x, \quad \dot{\phi}_{ts}(x, h) = \sum_{\ell=1}^m X_\ell(\phi_{ts}(x, h)) \dot{h}_t^\ell, \quad \dot{h}_t^\ell = \left\langle \lambda_s, \left(\frac{\partial}{\partial x} \phi_{ts}(\gamma_s, h) \right)^{-1} X_\ell(\phi_{ts}(\gamma_s, h)) \right\rangle.$$

On differentiating in λ_s in the direction η_s , we find that $v_t = (\partial/\partial h)\phi_{ts}(\gamma_s, h)k$, where k satisfies, for $t \in [s, 1]$,

$$\dot{k}_t^\ell = \langle \eta_s, u_{ts}^{-1} X_\ell(\gamma_t) \rangle + A_{ts}^\ell(\lambda, k). \quad (60)$$

Here, $u_{ts} = (\partial/\partial x)\phi_{ts}(\gamma_s, h) = u_t u_s^{-1}$ and

$$A_{ts}^\ell(\lambda, k) = \left\langle \lambda_s, \frac{\partial}{\partial h} \left[\left(\frac{\partial}{\partial x} \phi_{ts}(\gamma_s, h) \right)^{-1} X_\ell(\phi_{ts}(\gamma_s, h)) \right] k \right\rangle.$$

By writing a differential equation for $A_{ts}^\ell(\lambda, k)$, we find a constant $C < \infty$ such that, for all $t \in [s, 1]$,

$$|A_{ts}(\lambda, k)|^2 \leq C \int_s^t |\dot{k}_r|^2 dr.$$

It follows that the equation (60) uniquely determines k . Note that, if we set $v_t = 0$ and $\dot{k}_t = 0$ for $t \in [0, s]$, then we have, for all $t \in [0, 1]$,

$$v_t = \frac{\partial}{\partial h} \phi_t(x, h)k, \quad \dot{k}_t^\ell = \langle u_s^* \eta_s, u_t^{-1} X_\ell(\gamma_t) \rangle 1_{\{t \geq s\}} + A_{t0}^\ell(\lambda, k). \quad (61)$$

Proposition 8.1. *Let $\gamma \in H^{x,y}$ be tame, regular and minimal and let $v \in T_\gamma H^{x,y}$. Then $Q(v) = 0$ if and only if $v_t = J_t \eta_0$ for all $t \in [0, 1]$, for some $\eta_0 \in T_x^* M$.*

Proof. Choose a sub-Riemannian structure X for a . Suppose that $Q(v) = 0$. Set $k = k(\lambda, v)$, then $q(k) = 0$. On differentiating the identity (59) in h , in the direction k , we obtain

$$u_1^{-1} \frac{\partial^2}{\partial h^2} \phi_1(x, h)(k, k') = \sum_{\ell=1}^m \int_0^1 (\dot{k}'_t)^\ell \frac{\partial}{\partial h} \left[\left(\frac{\partial}{\partial x} \phi_t(x, h) \right)^{-1} X_\ell(\phi_t(x, h)) \right] k dt$$

so, by Proposition 7.5,

$$q(k, k') = \sum_{\ell=1}^m \int_0^1 (\dot{k}'_t)^\ell \left\{ \dot{k}_t^\ell - \left\langle \lambda_0, \frac{\partial}{\partial h} \left[\left(\frac{\partial}{\partial x} \phi_t(x, h) \right)^{-1} X_\ell(\phi_t(x, h)) \right] k \right\rangle \right\} dt. \quad (62)$$

Since γ is minimal, q is non-negative on K , so $q(k, k') = 0$ for all $k' \in K$. Hence there exists an $\eta_0 \in T_x^*M$ such that

$$\dot{k}_t^\ell = \langle \eta_0, u_t^{-1} X_\ell(\gamma_t) \rangle + A_{t0}^\ell(\lambda, k).$$

As we argued above, this forces $v_t = J_t \eta_0$ for all t . On the other hand, if $v_t = J_t \eta_0$ then, by the same calculations, we see that $v = (\partial/\partial h)\phi(x, h)k$ where k satisfies $q(k, k') = 0$ for all $k' \in K$. But $k \in K$, so $Q(v) \leq q(k, k) = 0$. \square

Proposition 8.2. *Let $\gamma \in H^{x,y}$ be tame, regular and minimal and let $\eta_0 \in T_x^*M$. Then $J_t \eta_0 = 0$ for all t only if $\eta_0 = 0$.*

Proof. Fix a sub-Riemannian structure X for a and a chart along γ . Write $\psi_t(\lambda_0 + \varepsilon \eta_0) = (\gamma_t^\varepsilon, p_t^\varepsilon)$ and $\dot{h}_t^{\varepsilon, \ell} = \langle p_t^\varepsilon, X_\ell(\gamma_t^\varepsilon) \rangle$. Set

$$v_t = J_t \eta_0 = \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} \gamma_t^\varepsilon, \quad r_t = \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} p_t^\varepsilon, \quad \dot{k}_t^\ell = \frac{\partial}{\partial \varepsilon} \Big|_{\varepsilon=0} \dot{h}_t^{\varepsilon, \ell}.$$

By differentiating the bicharacteristic equations we obtain, for all $t \in [0, 1]$,

$$\dot{v}_t = \sum_{\ell=1}^m \nabla X_\ell(\gamma_t) v_t \dot{h}_t^\ell + X_\ell(\gamma_t) \dot{k}_t^\ell, \quad \dot{k}_t^\ell = \langle r_t, X_\ell(\gamma_t) \rangle + \langle p_t, \nabla X_\ell(\gamma_t) v_t \rangle.$$

But $v_t = J_t \eta_0 = 0$ for all t so $\sum_{\ell=1}^m X_\ell(\gamma_t) \dot{k}_t^\ell = 0$ and so $|\dot{k}_t|^2 = \sum_{\ell=1}^m \langle r_t, X_\ell(\gamma_t) \rangle^2 = 0$ for all t . Now

$$\dot{h}_t^{\varepsilon, \ell} = \langle p_t^\varepsilon, X_\ell(\gamma_t^\varepsilon) \rangle = \langle \lambda_0 + \varepsilon \eta_0, (u_t^\varepsilon)^{-1} X_\ell(\gamma_t^\varepsilon) \rangle \quad (63)$$

where

$$\dot{u}_t^\varepsilon = \nabla X_\ell(\gamma_t^\varepsilon) u_t^\varepsilon \dot{h}_t^{\varepsilon, \ell}, \quad u_0 = \text{id}.$$

By differentiating this equation, we see that $(\partial/\partial \varepsilon)|_{\varepsilon=0} u_t^\varepsilon = 0$, so on differentiating (63) we obtain

$$\langle \eta_0, u_t^{-1} X_\ell(\gamma_t) \rangle = 0$$

for all t , and this implies that $\eta_0 = 0$ since γ is regular. \square

The next result is a characterization of the cut locus by path properties intrinsic to a . An analogous characterization in terms of a given sub-Riemannian structure is given in [6, Théorème 1.18]. We note the use of Proposition 8.2.

Theorem 8.3. *Let $x, y \in M$. The following are equivalent*

- (i) (x, y) lies outside the cut locus,
- (ii) there is a unique minimal path $\gamma \in H^{x,y}$, which is tame and regular, and the quadratic form Q is positive-definite on $T_\gamma H^{x,y}$.

Proof. In both (i) and (ii) we have a unique minimal path $\gamma \in H^{x,y}$. By Proposition 7.4, if γ is tame and regular then γ is the projection of a bicharacteristic, while any projected bicharacteristic is tame.

Suppose that Q is positive-definite and that $J_1\eta_0 = 0$ for some $\eta_0 \in T_x^*M$. Then $v_t = J_t\eta_0 \in T_\gamma H^{x,y}$ and $Q(v) = 0$, so $J_t\eta_0 = 0$ for all t , and so $\eta_0 = 0$, by Proposition 8.2. Hence J_1 is invertible.

Suppose, on the other hand, that J_1 is invertible. Since $\{v_1 : v \in T_\gamma H^{x,y}\}$ contains $\{J_1\eta_0 : \eta_0 \in T_x^*M\}$, we see that γ is regular. Suppose further that $Q(v) = 0$ for some $v \in T_\gamma H^{x,y}$. By Proposition 8.1, we have $v_t = J_t\eta_0$ for some $\eta_0 \in T_x^*M$. Then $J_1\eta_0 = 0$, so $\eta_0 = 0$, and so $v = 0$. Hence Q is positive-definite. \square

Recall from (15) that we define $K_t : T_y^*M \rightarrow T_{\gamma_t}M$ by

$$K_t\eta_1 = \left. \frac{\partial}{\partial \varepsilon} \right|_{\varepsilon=0} \pi\psi_{-(1-t)}(\lambda_1 + \varepsilon\eta_1)$$

and, for $s, t \in [0, 1]$ with $s \leq t$, set

$$C(s, t) = C(t, s)^* = J_s J_1^{-1} K_t^* \in T_{\gamma_s}M \otimes T_{\gamma_t}M.$$

Proposition 8.4. *Let $\gamma \in H^{x,y}$ be tame, regular and minimal. Suppose that x and y are non-conjugate along γ . Let $s \in [0, 1]$ and $\beta \in T_{\gamma_s}^*M$. For $t \in [0, 1]$, set $v_t^{\beta,s} = C(t, s)\beta$. Then $v^{\beta,s} \in T_\gamma H^{x,y}$ and, for all $v \in T_\gamma H^{x,y}$,*

$$Q(v, v^{\beta,s}) = \langle \beta, v_s \rangle.$$

Proof. Set $\eta_0 = J_1^{-1}J_{1s}\beta$. We can define $w = w^1 - w^2 \in T_\gamma H^{x,y}$ by setting $w_t^1 = J_t\eta_0$ for all t , and $w_t^2 = 0$ for $t \leq s$ and $w_t^2 = J_{ts}\beta$ for $t > s$. Then $w = (\partial/\partial h)\phi(x, h)k$, where $k = k^1 - k^2$ with

$$(\dot{k}_t^1)^\ell = \langle \eta_0, u_t^{-1}X_\ell(\gamma_t) \rangle + A_{0t}^\ell(\lambda, k^1), \quad (\dot{k}_t^2)^\ell = \langle u_s^*\beta, u_t^{-1}X_\ell(\gamma_t) \rangle 1_{\{t \geq s\}} + A_{0t}^\ell(\lambda, k^2).$$

Take $v' \in T_\gamma H^{x,y}$ and set $k' = k(\lambda, v')$. Then, using (62) for the second equality, we have

$$\begin{aligned} Q(w, v') &= q(k, k') = \sum_{\ell=1}^m \int_0^1 (\dot{k}'_t)^\ell \left\{ \dot{k}_t^\ell - A_{0t}^\ell(\lambda, k) \right\} dt \\ &= \sum_{\ell=1}^m \int_0^1 \langle \eta_0, u_t^{-1} X_\ell(\gamma_t) \rangle (\dot{k}'_t)^\ell dt - \int_s^1 \langle u_s^* \beta, u_t^{-1} X_\ell(\gamma_t) \rangle (\dot{k}'_t)^\ell dt = \langle \beta, v'_s \rangle. \end{aligned}$$

For $t \leq s$ we have

$$w_t = J_{t0} J_{10}^{-1} J_{1s} \beta = J_t J_1^{-1} K_s^* \beta = C(t, s) \beta.$$

Consider now the analogous construction in reverse time. Set $\eta_1 = K_0^{-1} K_{0s} \beta$ and define $\hat{w} = \hat{w}^1 - \hat{w}^2 \in T_\gamma H^{x,y}$ by setting $\hat{w}_t^1 = K_t \eta_1$, and $\hat{w}_t^2 = 0$ for $t \geq s$ and $\hat{w}_t^2 = K_{ts} \beta$ for $t < s$. Then, by the same argument, $Q(\hat{w}, v') = \langle \beta, v'_s \rangle$ for all $v' \in T_\gamma H^{x,y}$. But $w - \hat{w} \in T_\gamma H^{x,y}$, so this implies $Q(w - \hat{w}, w - \hat{w}) = 0$ and hence $\hat{w} = w$. So, for $t \geq s$, we have

$$w_t = \hat{w}_t = K_{t1} K_{01}^{-1} K_{0s} \beta = K_t (J_1^{-1})^* J_s^* \beta = C(s, t)^* \beta.$$

Hence $w = v^{\beta,s}$ and $v^{\beta,s}$ has the claimed property. \square

Theorem 8.5. *Let $x, y \in M$ and suppose that (x, y) lies outside the cut locus. Denote the unique minimal path in $H^{x,y}$ by γ . Recall the definitions (22) of the process W and (23) of the random variables $S(z)$ for $z \in \mathbb{R}^d$. Then $\mathbb{E}(e^{pS(z)/2}) < \infty$ for all $z \in \mathbb{R}^d$, for some $p > 1$.*

Define a new probability measure $\tilde{\mathbb{P}}$ on $\Omega^0(\mathbb{R}^m)$ by $d\tilde{\mathbb{P}}/d\mathbb{P} \propto e^{S(0)/2}$ and write μ_γ for the law on $T_\gamma \Omega^{x,y}$ of $Y(0) = v(W)$ under $\tilde{\mathbb{P}}$. Then μ_γ is a zero-mean Gaussian probability measure on $T_\gamma \Omega^{x,y}$ with the following properties

(i) *for all continuous linear functionals ϕ on $T_\gamma \Omega^{x,y}$,*

$$\int_{T_\gamma \Omega^{x,y}} \phi(v)^2 \mu_\gamma(dv) = Q(\tilde{\phi})$$

where $\tilde{\phi} \in T_\gamma H^{x,y}$ is given by $\phi(v) = Q(\tilde{\phi}, v)$ for all $v \in T_\gamma H^{x,y}$,

(ii) *for all $s, t \in [0, 1]$ with $s \leq t$,*

$$\int_{T_\gamma \Omega^{x,y}} v_s \otimes v_t \mu_\gamma(dv) = J_s J_1^{-1} K_t^*.$$

Moreover, the properties (i) and (ii) both characterize μ_γ uniquely.

Proof. Choose a sub-Riemannian structure X for a and consider the continuous quadratic form on $H^0(\mathbb{R}^m)$ given by

$$s(k) = \langle \lambda_1, (\partial/\partial h)^2 \phi_1(x, h)(k, k) \rangle.$$

Fix an orthonormal basis $(e_n : n \in \mathbb{N})$ for K which diagonalizes s on K . Formally, we have $S(z) = s(W(z))$ so, by a standard calculus for the Wiener chaos,

$$S(z) = \sum_n s(e_n) \langle W, e_n \rangle^2 + 2 \sum_n s(e_n, \tau(z)) \langle W, e_n \rangle + s(\tau(z)) \quad (64)$$

where the sums are understood in $L^2(\mathbb{P})$ and the random variables $\langle W, e_n \rangle$ are independent standard Gaussians. Set $s_n = -s(e_n)$ and $\sigma_n = s(e_n, \tau(z))$. We see from (64) that the series $\sum_n s_n$ converges and $\sum_n \sigma_n^2 < \infty$. In particular $s_n \rightarrow 0$ as $n \rightarrow \infty$. Since x and y are non-conjugate along γ , for all non-zero $k \in K$, we have

$$\|k\|^2 - s(k) = q(k) = Q(v(k)) + \|k'\|^2 > 0$$

where $k' = k - k(\lambda, v(k))$. Hence $1 + s_n > 0$ for all n . Denote by ν the standard Gaussian distribution on \mathbb{R} . Then, for all $s \in (-1, \infty)$ and all $\sigma \in \mathbb{R}$, we have

$$\int_{\mathbb{R}} e^{\sigma x} e^{-s x^2/2} \nu(dx) = \frac{1}{\sqrt{1+s}} e^{\sigma^2/(2(1+s))}. \quad (65)$$

Hence, for $p > 1$ sufficiently close to 1, for all $z \in \mathbb{R}^d$, we have

$$\mathbb{E}(e^{pS(z)/2}) = e^{ps(\tau(z))} \prod_n \int_{\mathbb{R}} e^{p\sigma_n x} e^{-ps_n x^2/2} \nu(dx) = e^{s(\tau(z))} \prod_n \frac{1}{\sqrt{1+ps_n}} e^{p^2 \sigma_n^2/(2(1+ps_n))} < \infty.$$

Let ϕ be a continuous linear functional on $T_\gamma \Omega^{x,y}$ and write $\tilde{\phi} = v(k^*)$, where $k^* = k(\lambda, \tilde{\phi})$. Then

$$\phi(v(e_n)) = Q(\tilde{\phi}, v(e_n)) = q(k^*, e_n) = \langle k^*, e_n \rangle q(e_n).$$

From (65), we see that under $\tilde{\mathbb{P}}$, the random variables $\langle W, e_n \rangle$ are independent zero-mean Gaussians of variance $(1 + s_n)^{-1} = q(e_n)^{-1}$. Now

$$\phi(Y) = \phi(v(W)) = \sum_n \langle W, e_n \rangle \phi(v(e_n))$$

so the law μ_γ of Y on $T_\gamma \Omega^{x,y}$ under $\tilde{\mathbb{P}}$ is a zero-mean Gaussian measure with

$$\int_{T_\gamma \Omega^{x,y}} \phi(v)^2 \mu_\gamma(dv) = \mathbb{E}(\phi(Y)^2) = \sum_n \phi(v(e_n))^2 / q(e_n) = \sum_n \langle k^*, e_n \rangle^2 q(e_n) = q(k^*) = Q(\tilde{\phi})$$

as claimed. Finally, by Proposition 8.4, for $\alpha \in T_{\gamma_s}^* M$ and $\beta \in T_{\gamma_t}^* M$, we have

$$\int_{T_\gamma \Omega^{x,y}} \langle \alpha, v_s \rangle \langle \beta, v_t \rangle \mu_\gamma(dv) = Q(v^{\alpha,s}, v^{\beta,t}) = \langle \alpha, C(s, t) \beta \rangle$$

so μ_γ also has the claimed covariance. The uniqueness statements are standard. \square

References

- [1] A. Agrachev, D. Barilari, and U. Boscain. Introduction to Riemannian and sub-Riemannian geometry. SISSA preprint, 2012.
- [2] A. A. Agrachëv. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.
- [3] N. Aida, S. Kusuoka, and D. Stroock. On the support of Wiener functionals. In K.D. Elworthy and N. Ikeda, editors, *Asymptotic problems in probability theory*, number 284 in Pitman research notes in mathematics, Sanda & Kyoto, 1990. Taniguchi international Symposium.
- [4] Robert Azencott. *Géodésiques et diffusions en temps petit*, volume 84 of *Astérisque*. Société Mathématique de France, Paris, 1981. Probability Seminar, University of Paris VII, Paris.
- [5] Davide Barilari, Ugo Boscain, and Robert W. Neel. Small-time heat kernel asymptotics at the sub-Riemannian cut locus. *J. Differential Geom.*, 92(3):373–416, 2012.
- [6] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. École Norm. Sup.*, 21:307–331, 1988.
- [7] J. M. Bismut. *Large Deviations and the Malliavin Calculus*, volume 45 of *Progress in Math.* Birkhäuser, Basel, 1984.
- [8] Jean-Michel Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander’s conditions. *Z. Wahrsch. Verw. Gebiete*, 56(4):469–505, 1981.
- [9] J. D. Deuschel, P. K. Friz, A. Jacquier, and S. Violante. Marginal density expansions for diffusions and stochastic volatility I: Theoretical foundations. *Comm. Pure Appl. Math.*, 67(1):40–82, 2014.
- [10] J. D. Deuschel, P. K. Friz, A. Jacquier, and S. Violante. Marginal density expansions for diffusions and stochastic volatility II: Applications. *Comm. Pure Appl. Math.*, 67(2):321–350, 2014.
- [11] Pei Hsu. Brownian bridges on Riemannian manifolds. *Probab. Theory Related Fields*, 84(1):103–118, 1990.
- [12] Pei Hsu. Heat kernel on noncomplete manifolds. *Indiana Univ. Math. J.*, 39(2):431–442, 1990.

- [13] Yuzuru Inahama. Large deviation principle of freidlin-wentzell type for pinned diffusion processes. *Trans. Amer. Math. Soc.* to appear.
- [14] David Jerison. The Poincaré inequality for vector fields satisfying Hörmander’s condition. *Duke Math. J.*, 53(2):503–523, 1986.
- [15] David Jerison and Antonio Sánchez-Calle. Subelliptic, second order differential operators. In *Complex analysis, III (College Park, Md., 1985–86)*, volume 1277 of *Lecture Notes in Math.*, pages 46–77. Springer, Berlin, 1987.
- [16] W. Klingenberg. *Riemannian Geometry*. De Gruyter, Berlin, 1982.
- [17] Rémi Léandre. Majoration en temps petit de la densité d’une diffusion dégénérée. *Probab. Theory Related Fields*, 74(2):289–294, 1987.
- [18] Rémi Léandre. Minoration en temps petit de la densité d’une diffusion dégénérée. *J. Funct. Anal.*, 74(2):399–414, 1987.
- [19] J. Lierl and L. Saloff-Coste. Parabolic Harnack inequality for time-dependent non-symmetric Dirichlet forms. [arxiv 1205.6493](#), 2012.
- [20] J. Marsden. *Applications of Global Analysis in Mathematical Physics*. Publish or Perish, Berkeley, 1974.
- [21] L. Mesnager. *Estimation en temps petit de densités conditionnelles dans des problèmes de filtrage non linéaire*. PhD thesis, Paris XI, Juillet 1996.
- [22] S.A. Molchanov. Diffusion processes and Riemannian geometry. *Russian Math. Surveys*, 30:1–63, 1975.
- [23] Alexander Nagel, Elias M. Stein, and Stephen Wainger. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.*, 155(1-2):103–147, 1985.
- [24] Jacques Neveu. *Processus aléatoires gaussiens*, volume 1968 of *Séminaire de Mathématiques Supérieures, No. 34 (Été)*. Les Presses de l’Université de Montréal, Montréal, Que., 1968.
- [25] James Norris. Simplified Malliavin calculus. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 101–130. Springer, Berlin, 1986.
- [26] James R. Norris. Heat kernel asymptotics and the distance function in Lipschitz Riemannian manifolds. *Acta Math.*, 179(1):79–103, 1997.

- [27] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.
- [28] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.*, 32(2):275–312, 1995.
- [29] S. R. S. Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20:431–455, 1967.